Universidad Autónoma de Madrid
Facultad de Ciencias
Departamento de Biología Molecular
Programa de Doctorado en Biociencias Moleculares

Parvovirus capsid modifications in cancer cell entry and evolutionary invasion of primate genomes

Nooshin Bayat

Director: José María Almendral del Rio

Centro de Biología Molecular Severo Ochoa (CSIC-UAM)

Madrid, 2018

This study was performed in the Center of Molecular Biology Severo Ochoa (CSIC-UAM). The PhD contract was financed by the Campus de Excelencia Internacional UAM + CSIC, which I would like to acknowledge the support.

I would like to dedicate my thesis to:

My father who is the most important person in my life,

and the memory of my beloved mother who passed away after a courageous battle with cancer while I was performing my PhD.

چنین گفت پیغمبر راستگوی

زگهواره تا گور دانش بجوی

فردوسی (۳۱۹ الی ۴۰۴ شمسی)

This is what the truthful prophet said
“Seek knowledge from the cradle to the grave”

Eso dijo el verdadero profeta
“Busca el conocimiento desde la cuna hasta la tumba”

Ferdowsi, Persian Poet (940–1025)

Acknowledgments

Hereby I would like to take a chance to mention and appreciate everyone who made it possible for me to do my thesis.

Special thanks to:

Dr. Jose Maria Almendral (Pepe): for accepting me as his PhD student and giving me the opportunity to come a long way from Iran to Spain and perform a PhD, also for his great support in everything, helping me to see the life from other perspective.

The "Campus de Excelencia Internacional, Universidad Autónoma de Madrid, Consejo Superior de Investigaciones Científicas": for funding my stay in Madrid for 4 years and the short stay (3 months) in Bern, Switzerland.

Pepa: por ser "la Mamma" una de las personas más importantes en mi vida a la que estoy muy agradecida de haber conocido; la que se preocupó por mí como si fuera mi madre de verdad y me enseñó el arte, la cultura y la ortografía española y me acompañó en las risas y las lágrimas, por eso la llamaría mi mama española.

Los miembros de laboratorio 224: Aroa: por su colaboración y la amistad especial, "la hermana española", Carlos Gallego: por su colaboración y ayuda en la parte de endógenos de la tesis y por su asesoramiento informático y por sus ánimos, Alberto: por sus consejos tan útiles y su simpatía, Rebeca y Luisa: por su energía positiva, buena compañía y ayudas técnicas, Tania: polo seu sorriso e as caras máxicas (y Ya ta! ¿Cómo que ya ta?), Jon: por ayudarme con los protocolos, Marta, Marcos y Carlos Domínguez: por la buena compañía y compartir las experiencias y Clara (la hija): por su energía positiva y las lecciones sobre los tejidos.

All the lab members of Dr. Ros lab in Bern, Switzerland: Carlos Ros: por acogerme en su laboratorio durante la estancia breve de doctorado, sus consejos científicos, confiar en mí y desarrollar mis habilidades, Raphael (Rafa): for the scientific collaborations and such a nice friendship that the distance couldn't tear it apart, Oli: for teaching me how to use the real time PCR and being so positive and cool, Mael: pour être une amie si proche et attentive, il semble que nous connaître depuis de nombreuses années, Andrea, Jan and Marcus: for sharing good moments.

Un agradecimiento especial a Belén Pérez y Alejandra Gámez por darme la oportunidad de aprender muchas técnicas nuevas de bioquímica en su laboratorio y a todos los miembros del laboratorio 220, por ayudar a integrarme en el CBMSO y hacerme sentir como en casa, Alfonso: por enseñarme las palabras españolas con dibujos especiales, tener la paciencia de escucharme y echarme una mano en cuanto organización, Sandra: por sus consejos tanto científicos como personales y por compartir los buenos momentos conmigo, Patricia: por enseñarme a utilizar el AKTA y ser una buena amiga, Celia: por enseñarme las normas de cultivo celular y darme la oportunidad de disfrutar de mi primera boda española y hacerme la primera tía iraní, Lorena: por mostrarme muchos sitios desconocidos de Madrid y ser tan maja, Ana: por compartir actividades deportivas, Irene, Ainhoa, Álvaro, Laura, Alejandro, Esmeralda, Raquel, Julia, Cyntia, Kristel (tía) y el resto: por los consejos, buenos momentos y dejarme utilizar la PCR (jajaja).

Departamento de acogida de estudiantes internacionales, especialmente Elena Puente e Isabel Sánchez Castillejo por ayudarme con el permiso de residencia en España.

Departamento de Biología y Facultad de Ciencia de UAM: Magdalena Sanz y Alejandro Rodríguez, Estrella, Juan Carlos Mingorance Delgado, Ignacio Tejero Toledano y Lola Vallejo.

Servicio de microscopía confocal, especialmente Maite y Carmen por la ayuda y enseñanza en utilizar los instrumentos ópticos.

Servicio de microscopía electrónica, especialmente Maite Rejas por su ayuda en sacar fotos tan maravillosas de los virus.

Rosa, por ser tan maja y traernos las risas que valen un mundo.

Servicio informático, Pedro (Pemau): por su ayuda en informática, enseñarme los trabalenguas y su compañía en adivinar el futuro ;), Ángel y Diego: por solucionar los problemas técnicos de ordenador.

My family, especially my father and Fariba for their love, support and education, who were always there for me and helped me face difficult challenges in my life.

My sister Sogol, the memory of my mother who brings happiness to my life, for being so strong and supportive.

My persian friends: Farnoosh: a real close friend, for sharing every moment since we have known each other and her positive way of thinking and problem solving supports in scientific and personal matters, Mahsa: my little sisy, Rashin: for being such a caring friend and the best mom ever, Farid: for being so positive and motivating, Maryam: the best organizer, Mehdi: who introduced Spanish education system to me, Sepideh: for such a nice and disconnecting trip to Marbella, Mitra: for her professional advices, Neda, Mana, Pegah, Sahar, Elham and Houra: for all the good moments, and Razi school friends: for reestablishment of such a nice and pure childhood friendship after 30 years.

Mis amigos de España: Los pinkies (mi familia española): Eva, Laura, Génesis, Leticia, Raquel, Rasa, Anna, Canto, Giovanni y Pablo por tener mucha paciencia conmigo y regalarme todos los buenos momentos (las fiestas, bailes y sonrisas) que han ayudado mucho a desconectarme del mundo del trabajo para poder lograr más eficacia, Los ICEX: Daniel, Ismael, Pablo, Ion, Jila, Miguel, Fernando, María, Rubén, Gacho, Guillermo, Cano, Javi, Irina y Fran, por cambiarme la vida, mejorar mi nivel de castellano, sus planes chulos y ayudarme a integrar en España y Célio: pelos seus conselhos de escrita e de encorajamento durante os últimos dias da tese.

Y

El amor que durante toda mi vida ha sido una inspiración y un motivo muy importante para poder avanzar y seguir adelante.

Summary

The ratio of physical particle to infectious virus, or virus specific infectivity (Si), is a central but unclear paradox in virus-host cell interactions. We found that a mutant at a threonine residue of the immunosuppressive strain of the parvovirus minute virus of mice (MVMi), here named MVMt, which resulted in increased Protected Data harbored a Si more than one hundred times higher than that of the MVMi. The MVMt and MVMi capsids differed from each other in biophysical responses to heat stress *in vitro*, including epitope configurations, genome uncoating and disassembly. The increased Si of MVMt seemed to rely on a favored endosomal traffic during virus entry into the transformed human kidney cell line (NB324K). Although no increased nuclear delivery of the MVMt genome could be demonstrated. This study points to Protected Data as a key unsuspected factor for regulating the efficiency of virus entry.

Following nuclear genome delivery, parvovirus infection may lead to its integration into the host cell genomes, a process that has allowed multiple RNA and DNA viruses to evolutionary colonize the germ line (endogenization) of their hosts. Endogenous parvoviruses and other ssDNA virus elements (ESVEs) have been recognized widespread in animal genomes, yet their distribution in primates is unknown. We report here the screening for significant ESVEs in human and non-human primate genomes available in databases. *In silico* evidences were supported by molecular amplification and sequencing of ESVEs and their flanking regions at the insertion sites in genomic DNA samples of primates of which some does not have the genomic sequence data available yet. The ESVEs that were originated from ancient or extant uncharacterized viruses highly homologous to members of the *Dependoparvovirus* genus of the *Parvoviridae* were named Simian Dependoparvovirus Elements (SDEs). Most SDEs resulted from four independent insertion events along primate evolution. For example, the genome of some members of the *Cercopithecidae* family (as papio, mandrill, or african green monkey) harbored single, or exceptionally double, SDEs homologous to AAV2. Many SDEs conserve the catalytic and structural domains of *Rep* and the capsid genes, although the length of sequences varied. These findings provide novel insights of the potential biological significance for primate evolution.

Resumen

La relación partícula física / virus infeccioso, o infectividad específica (IE), es una paradoja central pero poco clara en las interacciones virus-célula huésped. Encontramos que un mutante en un residuo de treonina de la cepa inmunosupresora del parvovirus diminuto de ratón (MVMi), llamado MVMt, que aumenta la Datos Protegidos alberga una IE más de cien veces mayor que la de la del MVMi. La cápsida del MVMt difirió de la de MVMi en las respuestas biofísicas al estrés por calor *in vitro*, incluída la configuración de epítopos, desencapsidación del genoma y desensamblaje. El aumento de IE del MVMt parece depender de un mejor tráfico endosomal durante la entrada del MVMt en células humanas transformadas (NB324K). Aunque no se pudo demostrar un aumento en la entrega nuclear de su genoma. Este estudio apunta a la Datos Protegidos como un factor insospechado pero clave en la regulación de la eficacia de la entrada de virus.

Tras la entrega del genoma en el núcleo, la infección por parvovirus puede conducir a la integración en el genoma de las células, un proceso que ha permitido que virus de ARN y ADN colonicen evolutivamente la línea germinal (endogenización) de sus hospedadores. Muchos parvovirus endógenos y otros elementos de virus ssDNA (ESVE) se han identificado ampliamente distribuidos en genomas de animales, aunque se desconoce su distribución en primates. Presentamos aquí el cribado de ESVE en el genoma de humanos y algunos primates no humanos disponibles en las bases de datos. Las evidencias *in silico* se apoyaron por amplificación molecular, y la secuenciación de las ESVE y sus regiones flanqueantes en los sitios de inserción, utilizando muestras de ADN genómico de primates, algunas de las cuales sin que la secuencia de su genoma esté accesible en bases de datos. Los ESVE identificadas, que se originaron a partir de virus antiguos o existentes en la actualidad, pero no caracterizados, son altamente homólogos a los miembros del género *Dependoparvovirus* de la familia *Parvoviridae*, y los hemos denominado Elementos Dependoparvovirus de Simio (SDE). La mayoría de las SDEs resultaron de cuatro eventos de inserciones independientes a lo largo de la evolución de los primates. Por ejemplo, el genoma de algunos miembros de la familia *Cercopithecidae* (como papio, mandrill, o mono verde africano) alberga SDE únicos, o excepcionalmente dobles, homólogos a AAV2. Muchas SDEs conservan los dominios catalíticos y estructurales de las proteínas Rep y Capsida, aunque las longitudes de las secuencias son diversas. Estos hallazgos proporcionan nuevos conocimientos que podrían tener un significado biológico para la evolución de los primates.

Abbreviations:

2Nt: VP2 N-terminal domain

AAV: Adeno-associated virus

BLAST: Basic Local Alignment Search Tool

CE: Cytopathic Effect

ClQ: Chloroquine

CPV: Canine Parvovirus

DMEM: Dulbecco's Modified Eagle's Medium

EPE: Endogenous Parvovirus Element

ESVE: Endogenous ssDNA virus elements

FPV: Feline panleukopenia virus

HAU: Hemagglutination unit, the minimum quantity of capsids that hemagglutinates in the serial dilution of capsids.

HPA: Hours Post Aphidicolin

HPI: Hours Post Infection

ICTV: International Committee on Taxonomy of Viruses

IE: Infectividad Especifica

MOI: Multiplicity of infection

MVM: Minute Virus of Mice

MVMi: lymphotropic strain of minute virus of mice

MVMt (T-mutant, MVMt-mutant): Mutant strain of MVMi in a Threonine residue

NCBI: National Center for Biotechnology Information

NLM: Nuclear Localization Motif

NP40: Nonidet P-40

NPC: Nuclear Pore Complex

NS proteins: Non-structural roteins

PCR: Polymerase Chain Reaction

PFU: Plaque Forming Unit. Number of particles per unit volume capable of forming plaques in determined amount of cells.

PtPV: Protoparvovirus

RBS: Ribosome Binding Site

SCID: severely combined immunodeficient

SDE: Simian Dependoparvovirus Elements

SDS: Sodium Dodecyl Sulfate

Si: Specific infectivity

SSC: Saline-sodium citrate

TAE: Tris Acetate EDTA buffer

TBS: Tris-buffered saline

TBT-T: Tris-buffered saline and tween-20.

vDNA-RF: Viral DNA Replicative Form

VLP: Virus Like Proteins

VP: Virus Proteins

Table of Contents

Index of tables

Index of figures

Introduction

Introduction

Section A.

1. General principles of virus entry.

Viruses recognize and attach to macromolecules of different natures, on the cell surface acting as receptors. In a productive infection, attachment to the suitable receptor allows viral particles to internalize and start the infection. Entry is the earliest process of the virus life cycle, a key step of the unique capacity of viruses to deliver their nucleic acid into the host cell.

There are several events and factors to be considered when studying viral entry, such as attachment, signaling, fusion with the plasma membrane, endocytosis, endosomes, penetration, intracellular transport and uncoating. Upon receptor interaction, the receptor-mediated endocytosis triggers the trafficking of virus across the endosome. The endosome is characterized by its mild acidic pH that descends from 6.5 to 5.5 in early to late endosomes, and finally it ranges from 5 to 4.5 when the endosome is fused with lysosomes forming endolysosomes where proteolytic degradation occurs. These changes of pH are mainly due to the activity of the ATP-dependent proton pumps present in the membrane of both endosomes and lysosomes (Diering and Numata, 2014; Hu *et al.*, 2015; Yamauchi and Helenius, 2013).

The trafficking across endocytic pathways is regulated by membrane associated proteins belonging to the Rab family GTPases. Each endocytic compartment has a unique Rab protein that could be used as specific marker (Agola *et al.*, 2011; Hutagalung and Novick, 2011). One of the most important processes in this study is the traffic of non-enveloped icosahedral viruses across the endosomal compartment leading to viral entry, which was best studied in Poliovirus (Hogle, 2002), Adenovirus (Luisoni and Greber, 2016) and Polyomavirus (Tsai and Qian, 2010). In these virus systems a diversity of interacting factors and molecular triggers are accessed by the incoming particles to progress across the endosomal compartments, while profound structural rearrangements and proteolytic cleavage of the capsid, by different pH dependent or independent processes, allow complexes of capsid protein-nucleic acid reach the cytosol, and for the karyophilic viruses enter the nucleus (Hogle, 2002; Luisoni and Greber, 2016; Tsai and Qian, 2010).

2. The *Parvoviridae*.

2.1. Taxonomic structure of the family

The *Parvoviridae* is a family of small, non-enveloped, icosahedral viruses with linear single-stranded DNA (ssDNA) genomes, infecting an exceptionally wide range of animal hosts, from invertebrates to mammals, including humans. Diseases caused by parvoviruses in their natural hosts vary in severity from subclinical to lethal infections, depending on the virus and host factors such as age, immune-competence and health conditions (Kailasan *et al.*, 2015). A condensed version of the taxonomic structure of the family is shown in Figure 1 (Cotmore *et al.*, 2014). The present study concerns mainly to the Adeno Associated Virus (AAV) of the *Dependoparvovirus* genus, and the Minute Virus of Mice (MVM) of the *Protoparvovirus* genus. *Protoparvovirus* (PtPV), is a genus of the Parvovirinae sub-family currently including the following species: *Rodent protoparvovirus 1* (H-1 parvovirus (H-1PV), Kilham rat virus (KRV), LuIII virus, minute virus of mice (MVM), mouse parvovirus (MPV), tumor virus X (TRX), rat minute virus (RMV)); *Rodent protoparvovirus 2* (rat parvovirus 1); *Carnivore protoparvovirus 1* (canine parvovirus (CPV) and feline panleukopenia parvovirus (FPV); *Primate protoparvovirus 1* (bufavirus (BuV)) and *Ungulate parvovirus 1* (porcine parvovirus (PPV).

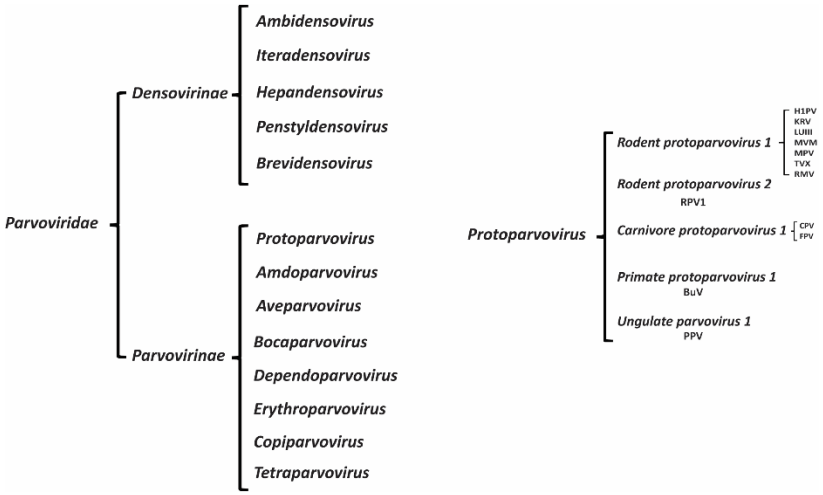


Figure 1. *Parvoviridae* taxonomic classification.
The subfamilies and genera of this family of viruses are displayed as proposed in the last version by the International Committee on Taxonomy of Viruses (ICTV), and the species belonging to *Protoparvovirus* genus are shown in the right.

2.2. Genome organization of the PtPV.

The PtPV ssDNA genome, of approximately 5 kb in length, is organized in two open overlapping transcriptional units (Figure 2), timely regulated (Cotmore and Tattersall, 1987; Tullis *et al.*, 1988). The left hand sequence driven by the P4 promoter encodes the NS1 and NS2 nonstructural proteins, and the right hand sequence driven by the P38 promoter encodes the VP1 and VP2 structural proteins. In spite of its small size, the parvovirus genome has maximized the coding capacity by alternative splicing, extensive post-translational modifications and proteolytic processing.

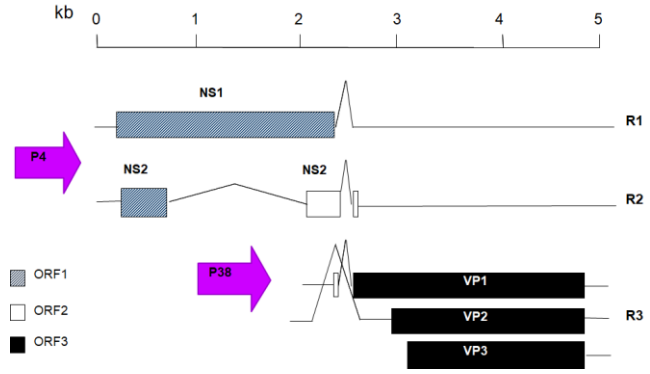


Figure 2. <u>Schematic view of the organization of MVM genome.</u>
The P4 promoter is used for NS1 and NS2 transcription and the P38 promoter is used for VP1, VP2 and VP3 transcription. The open reading frames (ORFs) are shown in boxes with different colors.

2.3. Functions of the parvovirus NS proteins.

Both NS polypeptides encoded by most PtPVs play multiple roles in virus life cycle. The smaller NS protein, NS2 protein (28 kDa), contains three isoforms arising from alternate splicing that can bind several cellular proteins and shuttle them from the nucleus to the cytoplasm via the Crm1 export pathway (Miller and Pintel, 2002). Functions assigned to NS2 include assisting capsid assembly (Cotmore *et al.*, 1997), messenger translation (Naeger *et al.*, 1993), DNA replication and virus production in a cell type specific manner (Naeger *et al.*, 1990). The larger NS protein, NS1 (82 kDa), is a multifunctional nuclear phosphoprotein, highly toxic for most cells, which performs crucial activities such as DNA-binding, Nickase, ATPase and helicase in the MVM unique rolling-hairpin mode of DNA synthesis (Cotmore and Tattersall*, 1995; Deleu *et al.*, 1997; Hörlein *et al.*, 2002; Miller and Pintel, 2002; Mouw and Pintel, 1998; Nüesch *et al.*, 1995; Nüesch and Tattersall, 1993).

2.4. The structure of parvovirus capsids.

The parvovirus capsid is a 25 nm-diameter assembly of 60 subunits in a *T* = 1 icosahedral symmetry formed by two to four largely overlapping proteins commonly designated VP1 to VP4 (Tijssen *et al.*, 2012). The 3-D atomic structure of PtPV capsid is available for many virus members of the family (Figure 3). Common features are the folding of their protein subunits into an eight-stranded antiparallel β-barrel topology, a β-cylindrical projection encircled by a canyon-like depression that surrounds the fivefold symmetry axes, and a pore at the center of the β-cylinder running between the surface and the interior of the capsid. The topology of the capsid surface differs among parvoviruses conferring characteristic virus host range. Flexible domains in the capsid are the N-terminal sequences of the VP subunits, which are dispensable for virus assembly, but strictly required for infection, as they contain functional motifs that can be hidden or exposed on the capsid surface at specific host cell compartments to mediate virus entry (Farr *et al.*, 2005; Lombardo *et al.*, 2002; Vihinen-Ranta *et al.*, 2002; Zádori *et al.*, 2001) and egress (Maroto *et al.*, 2004).

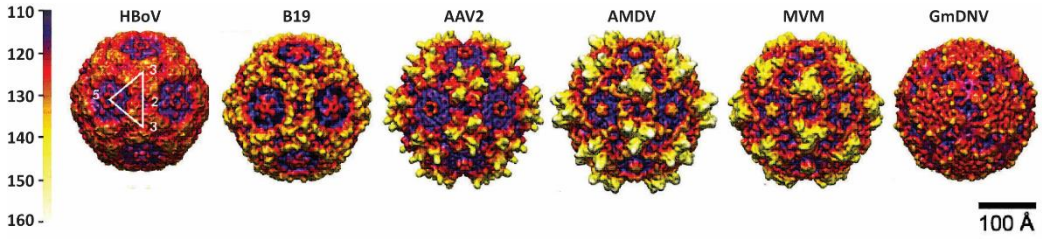


Figure 3. <u>Six representative parvovirus capsid structures.</u>
The cryo-reconstructions shown (7.9 Å resolution), with corresponding 2-fold views of B19, AAV2, AMDV, MVM, and GmDNV, highlight similarities and differences in the outer surfaces of the capsids. Vertical bar depicts color cueing as a function of particle radius in Å (from Gurda *et al.*, 2010)

3. The protoparvovirus life cycle.

It has been recently reported that the consecutive steps of the parvovirus life cycle (gene expression, nuclear translocation of proteins, capsid assembly, genome replication and encapsidation) are tightly coupled to the host cell cycle progression (Gil-Ranedo *et al.*, 2015). Importantly, parvovirus MVM gene expression in synchronous infections occurred at G1/S, implying that transcription does not require a previous viral DNA amplification, which indeed

occurred later at S/G2 phase. Capsid formation is particularly sensitive to cell cycle regulation, which is exerted at the level of non-conventional nuclear transport route(s) accessed by the assembly intermediates (Figure 4). Indeed, signals perturbing the cell cycle, as density arrest contacts or DNA synthesis stress, deregulate MVM nuclear capsid assembly.

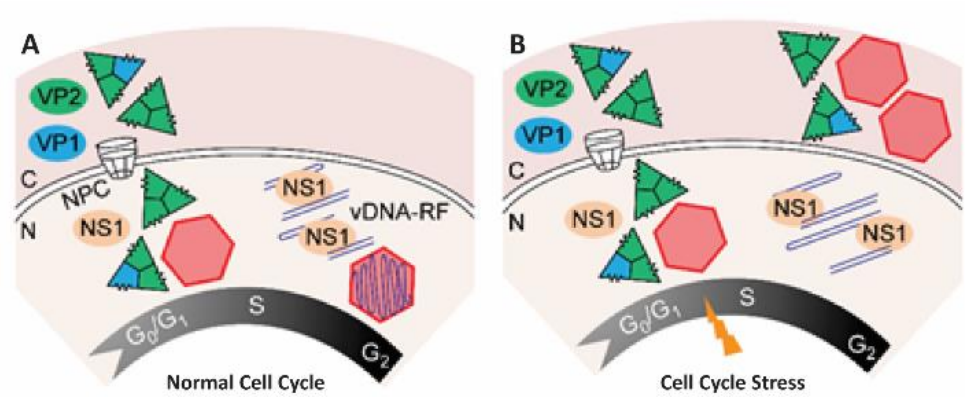


Figure 4. <u>The coupling of Parvovirus MVM and cell host life cycles.</u>
A: Parvovirus assembly and maturation is timely coupled to cell cycle progression from G1 to G2. Phosphorylated VPs trimers translocate into the nucleus at early S phase driven by the nuclear localization motif (NLM) (Lombardo *et al.*, 2000), and assemble into empty capsids. Viral genome replication () and virions maturation occur as cells become arrested at the late S/G2 phases by the infection. **B**: Infected cells subjected to cell cycle stress by thymidine DNA replication or density cell contacts (DA), interrupt drastically VPs nuclear transport, resulting into cytoplasmic empty capsids formation. C: cytoplasm, N: nucleus, NPC: nuclear pore complex and vDNA-RF: viral DNA replicative form. (taken from Gil-Ranedo *et al.*, 2015, with permission).

This coupling of parvovirus assembly to the cell cycle may largely rely on the control of capsid subunit phosphorylation. In MVM infection, the VP1 and VP2 structural subunits assembled into empty capsids were post-translationally modified through a complex and VP-specific pattern of phosphoserine and phosphothreonine residues (Maroto *et al.*, 2000). The 3-D structure of virus-like particles (VLPs) was obtained at high resolution (Hernando *et al.*, 2000), even though lacked phosphorylation (Riolobos *et al.*, 2010) and assembled in the cytoplasm of insect cells (Riolobos *et al.*, 2010; Yuan and Parrish, 2001), indicating that, at least for the VP2-only capsid, subunit phosphorylation is not important for icosahedral T=1 ordering.

However, nuclear transport of the VP2 homotrimer required cytoplasmic phosphorylation by the Raf-1 kinase (Riolobos *et al.*, 2010), although this phosphorylation was not sufficient to explain its cell cycle-regulated transport (Gil-Ranedo *et al.*, 2015). Raf-1 phosphorylation targets serine

residues of 2Nt (Maroto *et al.*, 2000), but localization of the many other phosphorylation sites in the VP1 and VP2 capsid subunits and their functions in the steps of the viral life cycle are unknown. It has been recently reported (Gil-Ranedo *et al.*, 2018) that the VP1 subunits are hyperphosphorylated in cytoplasmic assembly intermediates, but are subjected to an orchestrated dephosphorylation program during assembly that correlates with changes in Nt configuration (see Figure 5). These processes may help understanding virus/cell life cycles coupling. Furthermore, empty and DNA-filled virus particles drastically differed in the phosphorylation status of their VP1 and VP2 protein subunits, suggesting that genome encapsidation may imply VPs-dephosphorylation (Gil-Ranedo *et al.*, 2018).

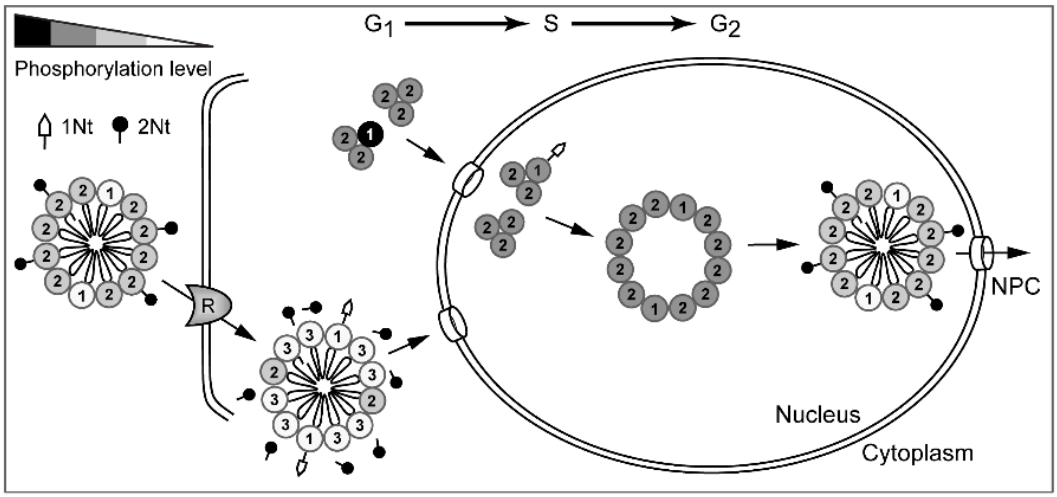


Figure 5. <u>Capsid subunits phosphorylation in MVM life cycle.</u>
The figure illustrates the differential phosphorylation levels and the exposure of the Nt-domains of the VP1, VP2 and VP3 structural proteins (respective numbers in this figure) in the complexes that traverse the cytoplasmic and nuclear membranes during the MVM infection cycle. The G_1, S, and G_2 letters (connected by arrows above in the figure) refer to the cell cycle steps at which virus assembly and maturation preferably occur. The VP protein subunits in the viral particles and assembly intermediates are depicted at their estimated stoichiometry. The VP1-Nt and VP2-Nt domains are only illustrated in exposed configuration(s) accessible to antibodies. R: Virus receptor; NPC: Nuclear pore complex. (taken from Gil-Ranedo et al., 2018, with permission).

4. Protoparvovirus entry into cells.

The PtPV entry, includes a diversity of processes such as virus-cell interactions, sequential exposure of structural protein domains and capsid rearrangements (reviewed in Cotmore and Tattersall, 2007; Harbison *et al.*, 2008; Parrish, 2010; Ros *et al.*, 2017; Vihinen-Ranta *et al.*, 2004). As summarized below, starting from the initial interaction with host cell receptors, the incoming

PtPv viral particle undergoes sequential steps leading to delivery of its ssDNA genome into the nucleus.

4.1. Receptors use.

Different types of biomolecules, such as proteins and carbohydrates, have been identified as attachment factors and in some cases as functional receptors for members of the PtPV genus. One of the best studied examples is the interaction of transferrin receptor (TfR) with CPV and FPV. These viruses use TfR to bind and infect cells (Parker *et al.*, 2001). Binding to TfR correlates with host specificity of these viruses, as FPV binds the feline TfR only, while CPV binds both the canine and feline TfRs. CPV variants accumulates changes on the top and side of the three-fold spike of the capsid surface, which confer more efficient use of the canine TfR receptor and reduce binding to the feline TfR (Hueffer *et al.*, 2003, 2004). The receptor residues involved in binding localize to the TfR apical domain and substitution of single amino acid may shift TfR binding ability to CPV or FPV (Palermo *et al.*, 2003, 2006). The TfR molecules asymmetrically associated in vitro with a few of the sixty icosahedral equivalent binding sites on the CPV capsid near the three-fold axis of symmetry (Figure 6). The TfR-CPV capsid interaction may be compromised by alteration of key residues of both the receptor and the virus capsid binding domain (Kaelber *et al.*, 2012), such as mutations at residues 206–380 of the TfR apical domain that reduce binding to CPV capsids. In particular, the N-linked glycosylation site at position 384 in the canine TfR provided resistance to the carnivore parvoviruses that were circulating prior to about 1975. Later a variant virus was emerged that overcame this block (Allison *et al.*, 2014). The use of TfR receptors by CPV and FPV variants illustrates the capacity of the PtPV to adapt to receptors determining host range.

The Sialic acid (SA; N-acetyl neuraminic acid) glycan is another well-characterized PtPV receptor and probably the most commonly used in nature. Glycans represent major cell surface carbohydrate components, thereby provide a vast collection of cellular attachment factors for many viruses (Olofsson and Bergström, 2005). Biochemical studies have shown SA to serve as a common primary attachment factor for several PtPV, although the difference between unspecific attachment and functional interaction (as a receptor leading to infection), has only been clarified in a few viral systems. SA was first suggested to play a role as MVM binding site to cell surfaces (Cotmore and Tattersall, 1987), in the range of 4–8 × 10^5 SA molecules per cell in established cell lines (Linser *et al.*, 1977; Spalholz and Tattersall, 1983) and primary host cells (Rubio *et al.*, 2005). It was subsequently demonstrated that SA acts as functional receptor for the infection of MVM (Lopez-Bueno *et al.*, 2006), PPV (Boisvert *et al.*, 2010) and H1-PV (Allaume *et al.*, 2012) PtPv.

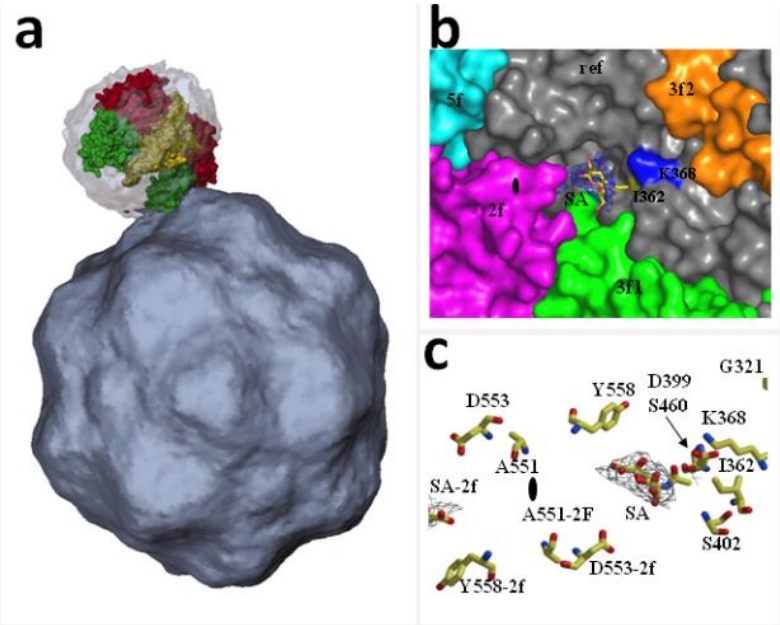


Figure 6. <u>Structures of receptor-protoparvovirus capsid complexes.</u>
A: Asymmetric binding of transferrin receptor to the CPV capsid. The image illustrates a TfR dimer placed manually into the cryoEM density of CPV capsid. The interaction is best resolved between one of the two apical domains of the TfR dimer (green) and the shoulder of a spike near a threefold icosahedral axis of CPV capsid. (Reproduced with permission from Hafenstein S. *et al.,* Proc Natl Acad Sci U S A. 2007 Apr 17;104(16):6585-9. Copyright (2007) National Academy of Sciences, U.S.A.). **B** and **C**: X-ray crystal structure identifying an infectious receptor attachment site on the parvovirus MVMp capsid soaked with sialic acid, showing the sugar electron density allocated in the twofold pocket (Reproduced with ASM permission from (Lopez-Bueno *et al.*, 2006).

4.2. *Protoparvovirus* uptake by the host cell.

Upon host receptor interactions, the PtPVs are internalized mainly by classical endocytosis (Clathrin-mediated endocytosis: CME). The endocytic route offers numerous advantages for the karyophilic parvoviruses. Endosomes provide a rapid and efficient transport for viruses, as the endolysosomal microenvironment with increasing acidic milieu and the changes in redox conditions, acid proteases and phosphatases enable capsids to undergo the structural conformational changes required for the infection, and eventually lead to the endosomal escape. But in addition to the classical CME, PtPV may use several alternative endocytic routes such as micropinocytosis (Boisvert *et al.*, 2010). Indeed MVMp may enter host cell by at least three potential endocytic routes (Garcin and Panté, 2015), mediated by clathrin, caveolin, and clathrin-independent carriers, the later restricted to transformed cells. Thus, both clathrin- and caveolin-mediated MVM endocytosis are dependent on dynamin in normal cells, but transformed cells

allow dynamin-independent mediated uptake of MVMp. These observations imply that although internalization by clathrin might represent the main entry route of PtPV, other alternative routes of uptake might be involved in parallel in certain cell types.

4.3. Endosomal traffic and capsid processing.

The endosomal processing of the incoming PtPV particles is essential for the infection. The endolysosomal microenvironment triggers crucial capsid structural rearrangements, and this process may be slow. MVM only reaches the cell nucleus after 8 hours post infection (hpi) when the DNA replication was detected and lysosomotropic drugs impaired the infection (Bowman *et al.*, 1988; Vihinen-Ranta *et al.*, 1998). For example, bafilomycin A_1, which raise the endosomal pH by inhibiting the vacuolar-type H^+ -adenosine triphosphatase (ATPase) (Hensens *et al.*, 1983; Ohkuma and Poole, 1978), or the weak base chloroquine diphosphate (that accumulates inside acidic compartment) inhibited MVM infection (Mani *et al.*, 2006). Endosomal acidification has been demonstrated to be essential for the infection of other parvoviruses (Basak and Turner, 1992).

The rearrangements that PtPV capsids undertake along entry can be traced by the configuration and functions of the N-terminal sequences of the VP1 and VP2 subunits. The amino termini of VP1 (VP1-Nt, 1Nt or VP1uR) and VP2 (VP2-Nt or 2Nt) are flexible domains whose configuration cannot be resolved in the X-ray crystal structure of PtPV capsids (Agbandje-McKenna *et al.*, 1998; Kontou *et al.*, 2005; Tsao *et al.*, 1991), and are normally concealed in their interior. These sequences carry essential transport signals and are thought to translocate across the five-fold channel of the capsid. This led to the suggestion that VP N-termini act as drivers of the traffic of intact virions across the cellular compartments along the virus life cycle (Maroto *et al.*, 2004). The externalization through the parvovirus fivefold capsid channel has been studied in detail for the VP2-Nt in virus like particles (VLPs) and native capsids (Castellanos *et al.*, 2013; Cotmore *et al.*, 1999, 2010; Farr *et al.*, 2006; Hernando *et al.*, 2000; Reguera *et al.*, 2004; Suikkanen *et al.*, 2003) . The dynamic exposure of Nt sequences across the channel is regulated by amino acid residues lying at the base of the pore, some of them configuring a hydrophobic gate tightly controlling the externalization of VP2-Nt and VP1 Nt sequences (Farr and Tattersall, 2004; Subramanian *et al.*, 2017). The VP2-Nt plays an elusive function during the virus entry process. This domain may be cleaved in vitro to form the VP3 protein by a chymotrypsin-like protease (Paradiso, 1981; Paradiso *et al.*, 1984; Tattersall *et al.*, 1977; Tullis *et al.*, 1992), and in vivo the VP2 to VP3 cleavage occurs in

the pH-dependent entry pathway, soon after PtPV internalization. Neither the resulting VP3 nor VP1 subunits can accomplish the essential entry function(s) harbored by the VP2-Nt domain (Sánchez-Martínez *et al.*, 2012). Therefore, it was proposed that the short VP2-Nt sequence may enlarge the functional diameter of the viral five-fold pore inside the endosome, determining the sequential exposure of VP2-Nt and VP1-Nt in the incoming metastable PtPV virion (Sánchez-Martínez *et al.*, 2012).

4.4. Endosomal escape and VP1 functions.

As other non-eneveloped viruses, the PtPV must penetrate, disrupt or breach, the host cell's delimiting membrane during entry, thereby undergo endosomal escape. But, the endosomal escape represents a major barrier for successful PtPV infection. Taking MVM as an example, a substantial proportion of the incoming MVM virions was demonstrated to follow a non-infectious pathway ending up in lysosomal compartments where they co-localized with lysosomal markers (Mani *et al.*, 2006). The PtPV endosomal escape depends on the profound capsid structural rearrangements occurring inside the endosome (see above), which lead to the externalization of the disordered VP1-Nt carrying a lipolytic activity (PLA2) that penetrates the endosomal membrane (Zádori *et al.*, 2001). The VP1-Nt becomes exposed in early endocytic vesicles (Cotmore and Tattersall, 1989; Farr *et al.*, 2005; Mani *et al.*, 2006). Complementation assays between wild-type and mutant MVM particles have been used to demonstrate the role of PLA_2 in mediating phospholipid bilayer penetration (Farr *et al.*, 2005). However, the precise site of PtPV endosomal escape is still incompletely understood. The optimal pH for the parvoviral PLA_2 enzymatic activity ranges between pH 6 to 7, but this activity drastically decreases at a pH below 5 (Canaan *et al.*, 2004). Additionally, brefeldin A, an antibiotic blocking the transition between early and late endosomes has been demonstrated to abrogate MVM infection (Ros *et al.*, 2002). These results suggest that only a minority of virions that enter the cytosol escape from an intermediate pre-lysosomal vesicle, namely late endosomes (Mani *et al.*, 2006).

4.5. Nuclear targeting and uncoating of the *Protoparvovirus* virion.

The transport of macromolecules through the nuclear pore complex (NPC) into the nucleus is a highly selective process mainly mediated by nuclear localization signals (NLS) exposed on the cargo molecules and recognized by specialized transport proteins (Timney *et al.*, 2016; Wente and Rout, 2010). The NPC central channel has a functional diameter up to a maximum of 39 nm, in the range of the diameter of the parvovirus capsids. Therefore, unlike larger viruses, which partially or

completely uncoat prior translocation through the NPC (Fay and Panté, 2015), the incoming parvovirus particle could pass through without disassembly. Two distinct mechanisms have been proposed for the nuclear entry of incoming PtPV: (i) PtPV particles translocate across the NPC mediated by exposed NLS in the VP1 unique region (Boisvert *et al.*, 2014; Lombardo *et al.*, 2002; Vihinen-Ranta *et al.*, 1997) and (ii) NPC-independent nuclear entry following local disintegration of the nuclear membrane (Cohen *et al.*, 2006).

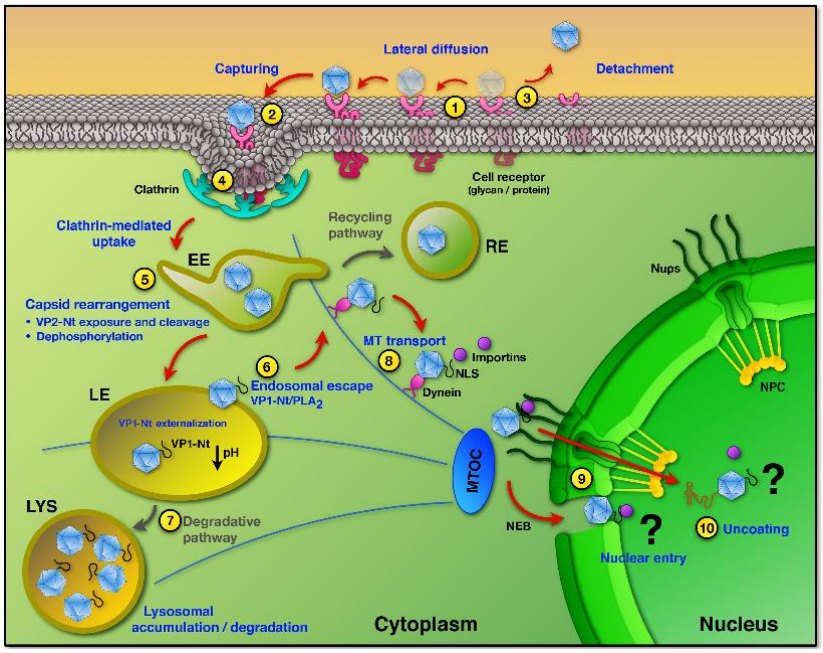


Figure 7. <u>Model of PtPV cell entry and nuclear import.</u>
(1) In the productive infection, PtPV capsids bind to one or more cell receptors with adequate affinity. (**2)** Following lateral diffusion, capsids are captured by pre-formed or forming clathrin pits (**3)** or otherwise become detached from their receptors. (**4)** Capsids are internalized via clathrin-mediated endocytosis (**5)** and enter the endocytic pathway where the low pH environment and local enzymes trigger structural rearrangements. (**6)** Major structural changes are a *de novo* VP2 N-terminus exposure and cleavage, and particle dephosphorylation, leading to VP1u exposure through the channel. The endosomal trafficking may vary in complexity and time, but commonly only a small proportion of particles escape after altering the endosomal membrane through the activity of the VP1-encoded phospholipase A_2. **(7)** Many incoming infectious particles fail to escape from endosomes and are routed to and accumulated in the degradative lysosomes. **(8)** In the cytosol, parvoviruses exploit the cytoskeleton and motor proteins to move to the nuclear vicinity. **(9)** It is not yet clear how the viruses and/or their genomes enter the nucleus, as nuclear pore-dependent entry and permeabilization of the nuclear envelope have been proposed. **(10)** The site and timing of capsid uncoating are also uncertain processes (Reproduced from Ros *et al.,* 2017).

In the later mechanism, the initial interaction of incoming PtPV capsids with NPC proteins would cause capsid structural rearrangements, which initiate a signal cascade resulting in local nuclear

envelope disintegration. These apparently contradictory observations may however function in concert to promote PtPV nuclear entry. An initial docking of the PtPV capsids to the NPC could trigger a local disintegration of the nuclear envelope, which would facilitate the transport of the capsids across the NPC. Finally, the mechanism and intracellular site of PtPV uncoating is not well understood (Figure 7).

5. The MVM (i,p) biology.

5.1. Cell hosts of MVMp and MVMi determining tropism *in vitro* and *in vivo.*

Tropism is a major pathogenic factor for infectious agents. Both host range variants of MVM, the lymphohematotropic (MVMi) and fibrotropic (MVMp) strains (Bonnard *et al.*, 1976; Crawford, 1966), provide a useful model for the study of DNA virus tropism in disease of a mammal host. MVM tropism is predominantly controlled by the gene encoding the 25nm-diameter T=1 icosahedral capsid (Ball-Goodrich and Tattersall, 1992; Maxwell *et al.*, 1995), that consist of two proteins, the VP2 (64 kDa), and VP1 (83 kDa) polypeptides. Although MVMp and MVMi differ only at few residues in VP2, they are reciprocally restricted for growth in each other's host cell type at a point after attachment to the cell surface receptor but prior to the onset of viral gene expression and DNA replication (Spalholz and Tattersall, 1983; Tattersall and Bratton, 1983). In established cell lines *in vitro,* two amino acid residues in MVMi, at VP2 codons 317 and 321, act coordinately as major determinants for the acquisition of fibrotropism (4), whereas the switch to lymphotropism for MVMp requires both an equivalent region of the MVMi coat protein gene and a segment of the non-structural (NS) protein genes (Hirt *et al.*, 1998).

In mice, MVMi is pathogenic for the newborn (Brownstein *et al.*, 1991; Ramírez *et al.*, 1996), and in severe combined immunodeficient (SCID) adults causes a depletion of hemopoietic precursors in bone marrow leading to severe leukopenia and accelerated erythropoiesis (Segovia *et al.*, 1999), which results from the ability of this virus to target committed, as well as long-term repopulating, hemopoietic stem cells (Segovia *et al.*, 1991, 2003). In contrast, MVMp is apathogenic in newborn mice (Kimsey *et al.*, 1986), and in oronasally-inoculated adult SCID mice its infection is asymptomatic with no clinical morbidity (Rubio *et al.*, 2005). However in intravenous inoculations MVMp evolves to pathogenic variants carrying single changes in the VP gene (Lopez-Bueno *et al.*, 2006; López-Bueno *et al.*, 2006).

5.2. Receptor use by the MVM strains.

The molecular basis of the distinct tropism of the MVM (p, i) strains has been a matter of intense debate. The interaction to the natural SA receptor used by both MVM strains should likely be involved with the tropism, and it has been consequently studied. Binding to an array of glycans enabled identification of SA types recognized by the two natural MVM strains, as well as by the emerged virulent capsid variants mentioned above (Nam *et al.*, 2006), and supported the evidences that lower affinity to SA glycans, rather than a different SA receptor specificity, determines MVM virulence (Lopez-Bueno *et al.).* However, it remains unclear how capsid binding to SA-glycans may determine MVM tropism in vitro and in mice. Tropism determinants of MVMp and MVMi capsids in vitro mapped to a few residues which also localized to the two-fold dimple (Antonietti *et al.*, 1988; Ball-Goodrich and Tattersall, 1992; Gardiner and Tattersall, 1988; Tattersall and Bratton, 1983). However, the N- and O- glycans and polar glycolipids composition of the cell lines used in those studies failed to explain MVM tropism solely based on the SA cell surface composition (Halder *et al.*, 2014), suggesting the requirement of other cellular factors. MVMp undergoes a remarkable switch of tropism from asymptomatic fibrotropic to pathogenic hematotropic behavior upon weeks post- infection in scid mice (Lopez-Bueno *et al.*, 2008). This dramatic change in pathogenicity involved the emergence of a heterogeneous viral population harboring multiple genetic changes, which were strictly confined to the surface of the dimple surrounding the SA-receptor binding pocket (Lopez-Bueno *et al.*, 2008). Therefore, the capsid residues mediating direct contact with SA are major determinants of MVM virulence and tropism along its evolution in mice. Further research will be required to elucidate whether these genetically heterogeneous variants co-operate to cause disease, use different SA molecules to infect distinct hematopoietic precursors, or arise in naturally infected host settings.

6. Biomedical applications of PtPv.

The biological relevance of the virus members of the *Parvoviridae* is not only due to their pathogenicity in their natural host, as two major biomedical applications can be outlined. Firstly, AAV2, the best study serotype of the *Dependoparvovirus*, is being developed as suitable vector for gene therapy of genetic diseases in human and animals e.g. Balazs *et al.*, 2012; Ran *et al.*, 2015. Secondly, AAV2 and some of the rodent PtPVs as H-1PV, MVM and LuIII, do display natural oncotropism and marked oncolytic activities against human cancer cells (Rubio *et al.*, 2001). The natural oncotropism of parvoviruses is determined, at least in part, by the fact that Parvovirus

replication is strongly dependent on host cell factors expressed during the S-phase, due to the limited coding capacity of the genome and their inability to induce resting cells to enter S-phase (Bashir *et al.*, 2000; Weitzman MD, 2006). Moreover, the virus cycle, particularly the nuclear translocation of structural subunits and capsid assembly, are tightly coupled to the cycle of the host cell (Gil-Ranedo et al., 2015). The dependency on poorly characterized transformation factors could be the reason of the high susceptibility of certain human tumor cells to Parvovirus induced killing (Angelova *et al.*, 2015; Nüesch *et al.*, 2012; Riolobos *et al.*, 2010; Ventoso *et al.*, 2010). Therefore, these features, the non-pathogenicity for humans, and the low prevalence of pre-existing antibodies, make parvoviruses promising tools to treat some human tumors (Marchini *et al.*, 2015).

7. The oncolytic MVMt.

MVM exhibits oncolytic potential (Rubio *et al.*, 2001). We have previously attempted to enhance MVM oncolysis by engineering the NS proteins with novel functional and biochemical activities. The NS1 protein of the immunsupressive strain MVMi was manipulated by site-directed mutagenesis at different residues that may control post-translational modifications of this large polypeptide. While most mutants yielded non-viable viruses or a neutral phenotype, a single mutant at a Threonine residue resulted in a viable virus with apparent high multiplication capacity in culture (Guerra et al, 2000). This MVMi mutant virus, here named MVMt, could be produced and purified at large scale and was therefore subjected to further analysis of infectivity and capsid properties as described in this work. The effects that the NS1 mutation brings to the biochemical and replicative properties of this protein are currently under investigation.

Section B

1. Endogenous Viruses

Viruses are common pathogenic infectious entities, but virus-host interactions may also develop as complex non-pathogenic patterns and symbiosis, an even in some cases the host may benefit from the virus infection. ((Jagdale and Joshi, 2018; Roossinck and Bazán, 2017). In some situations viruses may infect germ line cells of organisms along evolution, and become integrated and inherited as endogenous virus elements (EVEs) into their chromosomes (Feschotte and Gilbert, 2012) Most EVEs are RNA+ viruses of the *Retroviridae* (ERV) widely represented in vertebrates, including mammals and primates (Weiss and Stoye, 2013)). Human endogenous retroviruses (HERVs) have invaded primates and human genomes as several waves along phylogeny.

However, non-retrovirus RNA viruses can also persist as endogenous DNA elements (Zinkernagel *et al.*, 1997), and recently many families of RNA+, and RNA- viruses, have been identified widely represented as endogenous elements in multiple organisms (Crochu *et al.*, 2004; Horie *et al.*, 2010; Taylor *et al.*, 2010). Flaviviruses in invertebrates (Crochu *et al.*, 2004), Filoviruses (Taylor *et al.*, 2010) and Bornaviruses in human and mammalian genomes (Horie *et al.*, 2010), are noteworthy examples.

2. Endogenization mechanisms of RNA viruses.

The efficient Retrovirus endogenization as DNA elements is linked to the fact that they are the only animal viruses requiring integration into the host genome during their life cycle. As integrated provirus the Retrovirus genome may be activated, remain latent, or be recombined and deleted to become an endogeneous DNA element vertically transmitted. The mechanism of insertion of the RNA viruses, which normally don't integrate into the host chromosome in a productive life cycle, is generally unknown. But one exception is Bornavirus, as it was found to be catalyzed by the activity of LINE-1 retrotransposons that perform a reverse transcription of mRNAs into DNA elements in human cells (Horie *et al.*, 2010). However, some other bornavirus elements did not show the characteristic features of LINE-1 driven integration activity. Putative functions on host genome of these insertions are still undefined, and possible evolutionary roles in host protection against related viral infections have been proposed (Chuong *et al.*, 2016; Horie *et al.*, 2010).

3. Endogeneous DNA viruses

Viruses with DNA genomes from different families have also been identified colonizing the germ line of plants (Bejarano *et al.*, 1996), vertebrate and invertebrate animals (Arbuckle *et al.*, 2010), and human genomes (reviewed in (Katzourakis and Gifford, 2010)). Most of the identified endogenous DNA virus elements correspond to some of the currently known ssDNA virus families classified by the ICTV, which are common infectious entities in the three domains of life (Krupovic, 2013). Indeed most identified endogenous ssDNA virus elements (ESVEs) belong to the *Parvoviridae* (Arriagada and Gifford, 2014; Belyi *et al.*, 2010) (see Table 1). As described above, this family of viruses (Cotmore *et al.*, 2014) are common exogenous infectious entities in a wide range of vertebrate and invertebrate hosts, including primates and humans. But, in spite of the frequent presence of ESVEs in many eukaryotes, the information on primate ESVEs segments is limited to an AAV insertion in baboon (Katzourakis and Gifford, 2010) lacking molecular confirmation, and a fragment with far putative relationship to AAV in humans (Liu *et al.*, 2011).

Table 1. Endogenous Parvovirus Elements (EPEs) in animals.

Subfamily	Genus	Host
Parvovirinae	*Dependoparvovirus*	Bat (Pteropus alecto)
		Elephant *Loxodonta africana*
		Dolphin (*Tursiops truncatus*)
	Protoparvovirus	Rat (*Rattus norvegicus*)
		Guinea pig (*Cavia porcellus*)
		Common brushtail possum (*Trichosurus vulpécula*)
	Amdoparvovirus	Hyrax (*Procavia capensis*)
Densovirinae	*Densovirus*	Green peach aphid (*Myzus persicae*)
		Thorny Stick Insect (*Aretaon asperrimus*)
		Brown marmorated stink bug (*Halyomorpha halys*)

4. Mechanisms of ssDNA parvoviruses integration into host genomes.

The widespread presence of endogenous genetic elements belonging to parvoviruses in animal genomes may be related to the remarkable capacity of some members of this viral family, to be inserted in the genome of cells in culture. The *Dependoparvovirus* adeno-associated virus serotype 2 (AAV2) exhibits a site-specific insertion mechanism (Berns and Linden, 1995) into chromosome 19q (Kotin *et al.*, 1992; Samulski *et al.*, 1991), and at other sites (Hüser *et al.*, 2010; Janovitz *et al.*, 2013) of human cells, which is driven by the activity of the Rep non-structural viral protein.

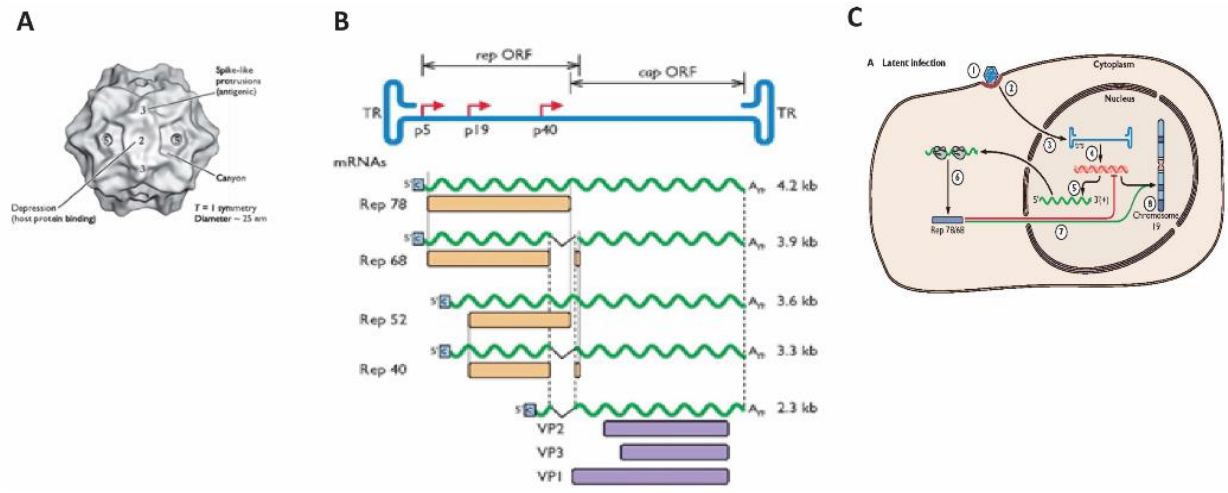


Figure 8. *Dependoparvovirus* AAV and insertion in the human genome.
A: Capsid features of AAV2. The topology of the spike, depression and canyon in the T=1 AAV2 capsid are illustrated. **B:** Genomic organization of AAV2, showing the three promoters and the distribution of the encoded Rep and Capsid proteins. **C:** Drawing of the latent infection of AAV2, in the absence of helper, leading to genome insertion in the human chromosome 19. (from Flint et al., 2008).

This unique feature made it an useful vector for several ongoing gene therapy applications (Buchholz et al., 2015). However, the capacity and site specificity of AAV seems to drastically differ from cultured cells to primary or natural cells. High prevalence of AAV unintegrated sequences are common in human tissues (Gao et al., 2004; Schnepp et al., 2005), and high number of infectious AAV circulate in monkeys (Gao et al., 2002, 2003). AAV2 may also insert its genome in other unrelated sites causing the activation of oncogenes that lead to severe diseases such as hepatocellular carcinoma (HCC) in mouse and humans (Donsante et al., 2007; Nault et al., 2015). On the other hand AAV5 may require different sequence organizations for its insertion (Janovitz et al., 2014), while other parvoviruses may become inserted at certain experimental conditions in cell culture (Hendrie et al., 2003).

Objectives

Objectives

The virus members of the *Parvoviridae* show wide distribution in nature causing inapparent or pathogenic infections in invertebrates and vertebrate animals including humans. Under certain interaction conditions they may integrate their genome into that of their host, and the integration may be stabilized into germline cells (endogenization). On the other hand, many species of these viruses are being extensively exploited as natural vehicles or genetically engineered vectors in gene therapy, and anti-cancer clinical trials in animals and humans. Despite these relevant biomedical and biotechnological properties, our knowledge on the molecular mechanisms mediating parvovirus cellular entry and endogenization in human and non-human primate genomes is still very limited. We therefore attempted to provide novel data on these issues addressing the following aims:

1. Extensive analysis of the chemical-physical properties of MVMi capsid, including thermostability and genome uncoating, compared with that of the MVMt (mutant) exhibiting a remarkably high specific infectivity.
2. Molecular analysis of the specific infectivity of these viruses (MVMi and MVMt) in the context of their entry pathways in human cancer cells.
3. Bioinformatic screening of primate genomes for the presence of parvovirus endogenous elements, confirming (whenever possible) their presence by molecular methods.
4. Characterization of the sequence and molecular structure of the confirmed endogenous primate parvovirus elements.

Materials

and

Methods

Materials and Methods

1. Cellular studies

1.1. Cell lines

The following cell lines were used in this study:

- NB324K (4K): Kidney fibroblasts of newborn human, transformed by SV40 (Shein and Eneders, 1962; Shein *et al.*, 1962). This cell line was kindly provided by B. Hirt, Epalinges, Switzerland.
- U87 MG: Astrocytes of human glioblastoma (ATCC: HTB14)
- Vero: Kidney epithelial cells of Cercopithecus aethiops (kindly provided by A. Carrascosa, CBMSO, Madrid)
- COS1: Kidney fibroblasts of Cercopithecus aethiops, transformed by SV40 (kindly provided by A. Carrascosa, CBMSO, Madrid)

1.2. Cell cultures

1.2.1. Materials

Mammalian cells were grown in Dulbecco's Modified Eagle's Medium (DMEM) containing 5% or 10% of Fetal Calf Serum (FCS). Cells were detached from the plastic petri dish with 25% Trypsin (Invitrogen), and stocks with minimal number of passages were preserved frozen in liquid nitrogen (see below). Cells were washed in Phosphate buffered saline (PBS), pH: 7.2-7.4, and inoculated with viruses in PBS supplemented with Mg^{++} and Ca^{++} (PBS^{++}). Cells were routinely maintained in an Incubator set at 37°C with 5% CO2 and humidified atmosphere, visually inspected by an inverted microscope (Leica Olympus), and the cell numbers were determined in Neubauer chamber (or hemocytometer).

1.2.2. Cell passages.

In order to provide nutrition, and remove the metabolic toxins, cells were passaged prior reaching full confluence. For this, the medium was removed by vacuum, the cell monolayers were washed with PBS, and trypsin (800-1000 µl/P100) was added. When significant number of rounded cells

was observed under microscope, the excess of Trypsin was removed and the cells were kept in the incubator for 1-3 minutes. When cells became completely detached, they were resuspended in the corresponding medium and 10 µl of the suspension was taken apart to be counted in Neubauer chamber. The required amount of cells for each seeding was calculated and added to each plate containing the growth medium.

1.2.3. Cells preservation.

In order to have fresh cells, and avoid possible contaminations and extensive passages, each cell line was frozen in aliquots in liquid nitrogen until needed. For this, a plate containing fresh cells growing in good conditions was washed with PBS, detached with trypsin as above, and resuspended in 2-3 ml of DMEM 10% media. The suspension was moved to a 15ml falcon tube, centrifuged at 1200rpm for 5 minutes (Sorvall). The cell pellet was resuspended in 1:1 ratio of a cocktail containing M-I medium (DMEM and 20% FCS) and M-II medium (DMEM and 20% FCS and 30% Glycerol) and it was aliquoted in cryotubes (Nunc). The aliquots were first kept in freezing container rack inside the -70°C freezer for at least 24h, afterwards the samples were kept in liquid nitrogen (-180°C – 210°C). For further use, frozen cells were incubated at 37 °C in a water bath until defrosted, then added to a falcon (F15) tube containing 10 ml of DMEM supplemented with 5% of FCS and centrifuged 5 min at 1500 rpm (Sorvall). The supernatant was removed, and the pellet was resuspended in the corresponding medium and finally seeded in appropriate plates (plastic petri dishes).

1.2.4. Cells infection with viruses.

One day prior to infection, the cells were seeded usually at a density of 2.5 x 10^5 cells per P60mm petri dish. For infection, cells were washed with PBS, the calculated amount of virus, diluted in 500 µl (per P60mm) of PBS^{++} with 0.1% of FCS, was added to the cells. A negative control (Mock), containing cells treated with the same amount of buffer but no virus, was considered for each experiment. After inoculation, the plates were shaken gently for one hour at 37°C, then the virus was removed and fresh medium was added. Assuming this time as zero point of infection (T0), the plates were kept at 37 °C inside incubator for the desired time.

2. Virus studies

2.1. Viral strains.

The Protoparvovirus Minute Virus of Mice (MVM) was used in all the experiments of this study. Viral strains were:

- MVMi (also referred as MVMi-wt), the immunosuppressive strain (Bonnard *et al.*, 1976; Segovia *et al.*, 1999) (Kindly provided by B. Hirt, Epalinges, Switzerland)

- MVMt mutant (also referred as T-Mutant), a Thr/Ala point mutant constructed in the non-structural (NS) gene of the MVMi genome (Guerra, 2000).

2.2. Virus Production and Purification

MVMi wt and mutant viruses were produced and highly purified following previously published methods (Gil-Ranedo *et al.*, 2015; Rubio *et al.*, 2001; Santarén *et al.*, 1993) with some modifications. Viruses were prepared from their respective molecular clones, pMVMi and pMVMt mutant (Guerra, 2000). These plasmids were previously amplified in JC8111 *Escherichia coli* strain, and transfected into NB324K human cells by electroporation. Cells were cultured 48 hours, lysed by freezing-thawing and clarified, and the intracellular virus was harvested and kept at -70 °C as reference stock. For large scale virus production, NB324K cells seeded in four P100 plates and grown to 80% confluence were infected with the reference virus stocks at low MOI (10^{-3} PFU/cell) and kept at 37°C in the incubator. After 7 days, the cells (showing extensive cytopathic effect), were scraped in the same medium and centrifuged at 5000 rpm for 15 minutes (4°C). The virus in the supernatant was recovered by adding 3.4% Polyethylene glycol (PEG) and 0.5M NaCl followed by 24h incubation at 4 °C and then 30 minutes of 5000 rpm centrifugation. The cell pellet was resuspended in PBS + 0.1% Sarkosil, dispersed by sonication at 20 kHz (2 min) and finally centrifuged at 800 rpm for 10 minute (4°C). The supernatants were next mixed and centrifuged through a 20% sucrose cushion+ 0.1% Sarkosil in a TST.28.38 rotor at 18krpm for 16h at 12 °C. Virus pellets were resuspended in PBS + 0.1% Sarkosil and subjected to Caesium Chloride (CsCl) equilibrium centrifugation adjusted to the density of 1.38 g/ml in a final volume of 10 ml. Samples were centrifuged at 50,000 rpm, 24h at 8 °C in a TFT 80.13 rotor. Fractions of 500µl were taken up to down into 1.5ml Eppendorf tubes and subjected to the hemagglutination assay (refer to M&M

2.3.1). The virus containing fractions (density range from 1.39 to 1.43 g/ml) were pooled and rebanded at the same conditions as above, to remove empty capsids. Finally, the fractions of the second gradient with hemagglutination activity and density corresponding to DNA-filled virus were pooled, and the excess of CsCl was removed by extensive dialysis against PBS. Viral stocks were kept in aliquots at -70°C, and maintained at 4°C for routine use within two weeks.

2.3. Virus titration

In order to have a reliable comparison between viruses, different viral titration methods were used.

2.3.1. Hemagglutination Assay (HA)

This method was used to determine the amount of virus particles (including DNA-filled virions and empty capsids) by their capacity to agglutinate mouse red blood cells. 1ml of heparinized mouse fresh blood was taken to 10ml in PBS (10 times of the blood volume) on ice and washed by centrifugation at 2000 rpm for 5 minutes and re-suspension in 10 ml of PBS (three times). A 96U well plate was filled with 50 µl of PBS in each well (98µl in the first row), then 2µl of capsid or virus sample (sample volume) was added to the first row (final volume 100 µl). Serial ½ dilutions starting with 50µl of the first row were made and 50µl excess of the last dilution was removed. Then 50µl of 1% recently washed red blood cells was added to each well and the plate was kept at 4°C for at least two hours. Finally, the hemagglutinated wells were counted (n) and viral particles were determined as hemagglutination units per each ml according to the following formula:

$$Hemagglutinating\ units\ (HAU) = \frac{2^n}{Sample\ volume}$$

2.3.2. Real Time PCR

This method was used to measure the amount of viral DNA by specific MVM primers (Ros *et al.*, 2006). It was set up and carried out during the short stay under supervision of Dr. Carlos Ros (Department of Chemistry and Biochemistry, University of Bern, Switzerland).

The standard DNA quantity was prepared by MVM-plasmid with known concentration (previously measured by A_{260} absorbance). Different dilutions of standard were made to normalize the measurement, and viral concentration was measured according to the standard.

The master mix was prepared as shown in Table 2. Then, 2 µl of DNA including various virus dilutions (10^{-1}, 10^{-2}, 10^{-3} and 10^{-4}) or standard dilutions (10^{-7}, 10^{-6}, 10^{-5} and 10^{-4}) were added to 18 µl of master mix. The PCR program (Table 3) was run on CFX96 (BioRad) and the viral amounts were estimated using the CFX Manager™ software.

Table 2. Components used for each reaction of Real Time PCR

Component	Details	Vol. (µl)	Final Con.
iTaq Supermix (2X)	Bio-Rad	10	1X
Primer Forward (10pM)	5'-GACGCACAGAAAGAGAGTAACCAA-3'	1	0.5 pM
Primer Reverse(10pM)	5'-CCAACCATCTGCTCCAGTAAACAT-3'	1	0.5 pM
ddH2O DNase- RNase free	Gibco	6	

Table 3. PCR program applied to quantify MVM DNA copies

Initial denaturation	95°C→ 300 s	
Denaturation	95°C→ 15 s 61°C→ 15 s 72°C→ 15 s	40 cycles
Annealing		
Extension		
Final denaturation	95°C→ 60 s	
Melting curve	65°C up to 95°C→ 0.1°C/s	

2.3.3. Electron Microscopy

The number of viral particles was also determined by a transmission electron microscope [JEOL (JEM1010)] operating from 12k to 100k magnitude. In order to visualize the viral particles, the negative staining with Uranyl Acetate was applied on 1- 10 µl of viral serial dilutions. After the ionization, the copper grids were put on the samples (facing the samples) for 3 minutes, then washed 3 times in ddH2O (facing the water) and kept on Uranyl Acetate less than 1 minute. Then samples were dried on filter paper for 5 minutes and eventually observed at 12k, 40k and 60k of magnitude. The amount of virus particles in 12K magnitude was counted and used for relative study of particles in viral stocks.

2.3.4. Infectivity Methods.

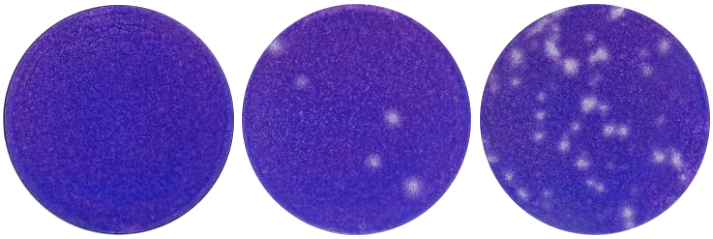
2.3.4.1. Plaque Assay (PA)

This assay was used to measure the number of viral particles that can form plaques in semi-solid medium. First, cells seeded on P60 or P35 dishes (the day before) were infected with various dilutions (e.g. 10^{-2}, 10^{-3}, 10^{-4} and etc.) of virus in 500µl of PBS^{++} (in case of P60 plates) for one hour at 37 °C. The virus inoculum was removed and a total of 7 ml of semi-solid medium (DMEM+antibiotics, 10% FCS, 0,8 % Agarose FMC SeaPlaque™) was added to each P60 plate. The plates were kept at 4°C for 5 minutes until the gel of Agarose was formed, and then placed at 37°C in the incubator. After 6 days, cells were fixed by adding 10% Formaldehyde (for at least 2 hours), stained with 0.2% Cristal Violet in 10% Formaldehyde (1 hour), and finally washed with ddH2O and dried. The number of plaque forming units was routinely determined in triplicates.

An example with serial dilutions is shown. The plaques were counted and then the concentration of Plaque Forming Unit (PFU)/ml was calculated according to the following formula

- D= dilution and V= Volume

$$\frac{PFU}{ml} = \frac{\text{Number of Plaques}}{\text{D x V}}$$

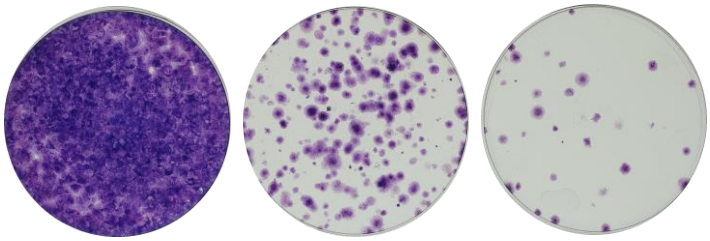


Mock 10^{-2} 10^{-1}

2.3.4.2. Colony Forming Assay (CFA)

The CFA was used to study cell survival after infection (based on Ramírez *et al.*, 1995). Two days prior to the experiments, cells were seeded in P60 plates to obtain 80% confluence. In the day of infection, cells were washed with PBS, then Trypsin was added for 2 minutes and immediately stopped with serum containing medium. The detached cells were collected and centrifuged at 2500rpm for 5 minutes, resuspended in 10ml of FCS containing medium, and then incubated at 37°C incubator (to regenerate the receptors). After 1 hour the cells were centrifuged at 2500rpm

for 5 minutes, resuspended in 1ml of medium, and the number of cells was counted in the hemocytometer (Neubauer chamber). For each infection, 10^5 cells were added to 1.5 ml Eppendorf tubes, serial amounts of virus were added to the cells (in 50µl of volume per each infection) and the infection was incubated at 37°C in a Thermomixer (Eppendorf). After one hour the cells were seeded in a serial dilution of 10^4, $10^3$and 10^2 cells/P60 dish and kept at 37°C in the incubator. After 6 hours a neutralizing anti-MVM antiserum was added to stop putative virus progression. After 10 days, the medium was removed, cells were washed with PBS, fixed with 4% formaldehyde (60-120 minutes), stained with Cristal Violet (1 hour), and washed with ddH2O. Finally, the cell colonies were counted and the number of survived cells/ml was estimated and used for relative study between both viruses (see illustrative example below).



Mock 10^{-2} 10^{-1}

2.4. Analysis of *in vivo* ^{32}P-radiolabeled capsid proteins.

^{32}P-labeled capsid subunits of MVM were produced and purified in NB324K cells following a previously described method (Gil-Ranedo *et al.*, 2018; Maroto *et al.*, 2000, 2004) with some modifications. Cells were infected at MOI 10, starved for 4 hours in phosphate-free DMEM supplemented with 5 % dialyzed FCS, and labeled from 4 to 24 hpi in the same medium containing 0.1 mCi per ml of [^{32}P] orthophosphate carrier free (Perkinelmer). Cells were washed and scraped into PBS supplemented with 0.2 % sodium dodecyl sulfate (SDS), protease inhibitors (1mM phenylmethylsulfonyl fluoride (PMSF); 10 µg/ml aprotinin; 10 µg/ml pepstatin; 10 µg/ml leupeptine) and phosphatase inhibitors (5 mM NaF, 20 mM β-glycerophosphate). The homogenates were boiled to disassemble capsids and denature VPs proteins, cooled to 4 °C, and debris were removed by low speed centrifugation (10,000 x g, 15 min). The supernatants were collected avoiding pellet contamination, and anti-VPs antibody (1/200) was added and incubated at 4 °C overnight under gentle rotation. Homogenates were next incubated with 10% protein A-sepharose equilibrated in binding buffer (BF: 50mM Tris-ClH, pH 7,5; 150 mM NaCl; 1mM EDTA, 1% NP-40, 0.1 %BSA) for 2h under shaking at 4 °C, and the immune-complexes were finally

recovered by low speed centrifugation (10,000 x g, 10 min). Pellets were washed (resuspension/incubation at 4 °C in BF/centrifugation) twice more before elution by boiling in Laemmli buffer at 95 °C for ten minutes. The ^{32}P-radiolabled VPs proteins were resolved by SDS-PAGE and visualized by autoradiography.

2.5. Enzymatic analysis of virus phosphorylation

Highly purified viral particles were subjected to a concentration gradient of lambda phosphatase enzyme (NE Biolabs). The reactions were prepared according to manufacturer's guide and also as described by Wolfisberg *et al.*, 2016.

3. Cellular analyses

Different types of microscopes were used to visualize the virus particles, or virus interaction with cells (Table 4).

Table 4. Microscopes used in this study

Type	Device Name	Manufacturer	Labeling
Florescent Microscope	Axiovert	Zeiss	FITC, Texas Red, DAPI
	CCD chamber	Zeiss	FITC, Texas Red, CY5, DAPI
	Multiphoton Confocal	Zeiss	FITC, Texas Red, CY5, DAPI
(TEM)	Electron Microscope	JEOL (JEM1010)]	Uranyl Acetate contrast

3.1. Immunofluorescence

3.1.1. Slide preparation

The round glass cover slips were submerged in ethanol 100% for 10 minutes. In each P60 plate, 6 cover glasses were adhered and dried under sterile air flow, the calculated amount of cells was seeded in each P60 plate, then incubated at 37 °C in the incubator. When the cells were 80% confluent the infection was performed (refer to M&M: 1.2.4.), afterwards the cells were washed with 2ml of PBS and fixed by 3ml of 4% Paraformaldehyde (PFA) for 8 minutes. The PFA was replaced with 3ml of PBS and the plates were kept at 4 °C until staining.

3.1.2. Fluorescent antibody staining

Each coverslip was taken to individual well of a M24 plate previously filled with PBS. Then the PBS was replaced by permeabilization buffer containing 500 µl of PBS and 0.1% Triton X-100, and samples were shaken gently for 10 minutes on rotator. The permeabilization buffer was replaced by 500 µl of PBS+0.1% Triton X-100 containing 1% FCS (blocking solution) for 20 minutes. After blocking, a 10 cm x 10cm Parafilm was placed on a filter paper in a humid box, and 25 µl of primary antibody was placed on Parafilm for each coverslip (see table 5 for dilutions). Covers were kept on Parafilm facing toward the antibody for 1 hour at 37 °C, then moved to a P24-dish (facing up) and washed 4 times with PBS (5 minutes on rotator each time). Then, the secondary antibody was applied in the same way as primary antibody, and samples were incubated at 37°C for 45-60 minutes. Eventually, the secondary antibody was washed by submerging the cover glasses 20 times in PBS and 2 times in ddH2O. Finally, 10 µl of MOVIOL+DAPI (Fluoromount) was placed on microscope slide and the coverslips were mounted on the slides placing the cells facing toward the Fluoromount. Slides were dried for 20 hours at room temperature, or 1 hour at 37 °C, and then kept at 4 °C until microscopic analysis.

Table 5. Antibody dilutions used for Immunofluorescence

Name	Source	Description	Type	Working dilution	Fluorophor emission
Anti-NS1	Rabbit	Non-structural NS1 protein of MVM	Primary	1/100	
Anti-Vps	Rabbit	denatured MVM capsid proteins	Primary	1/200	
Anti-B7	Mouse (Mab)	B7 epitope of MVM 3x Capsid	Primary	1/50	
Anti- MVM Capsid	Rabbit	Entire Capsid	Primary	1/200	
Anti Rabbit	Goat	FITC	Secondary	1/500	488nm
Anti Mouse	Goat	Texas-Red	Secondary	1/500	564nm

3.1.3. Fluorescent DNA in situ hybridization (FISH)

Cells fixed with PFA were permeabilized for 10 minutes by 500µl of PBS + 0.2% Triton X-100, then the samples were equilibrated with Saline-Sodium Citrate (SSC) 2X + 15% Formamide for 5 minutes. A humid box was prepared and then 10 ng of each oligo (labeled with CY5-alexa 647 nm) in 50 µl of FISH buffer (Gil-Ranedo *et al.*, 2015) was spotted on coverslips. Samples were incubated

in humid box at 37°C for 3 hours and finally washed 2 times with SSC 2X + 15% Formamide for 20 minutes at 37 °C. This method was generally followed by antibody staining, and observation under fluorescence microscopy (refer to M&M: 3.1.2).

4. Molecular analyses

4.1. Nucleic acids Extraction

4.1.1. Virus DNA extraction

The DNeasy Blood & Tissue Kit (Qiagen) was used, following the manufacturer's protocol, to extract the viral DNA to be used for real time PCR.

4.1.2. Tissue DNA extraction

Frozen tissue samples were sliced, and per each 10 g wet weight 700 µl of 50mM Tris-HCl (pH:8), 100mM EDTA (pH:8), 100mM NaCl and 1% SDS was added. The slides were incubated at 55 °C overnight with 18 µl of Proteinase K (20 mg/ml), then 4 µl of RNase A+T1 (Qiagen) was added and samples were incubated 30 minutes at room temperature. Afterwards, 250 µl of saturated NaCl (more than 5M) was added to samples, shaken, and then centrifuged at 14000 rpm for 5 minutes. 500µl of Isopropanol 100% was added to 850µl of supernatant, mixed well and centrifuged at 12000rpm for 5 minutes. The pellet was washed with 500-1000 µl of 75% Ethanol at 4 °C, centrifuged at 12000 rpm for 3 minutes in the cold room and air dried. Finally, the DNA was resuspended in 100µl of TE buffer (10mM Tris-HCl, 1 mM EDTA, pH 8).

4.2. Amplification of genetic material

4.2.1. Primer Design

The genomic DNA sequences were obtained from GenBank (www.ncbi.com/genbank). The Generunner software (Ver. 05.0.63) was used to select the best oligos according to their GC content, annealing temperature, secondary structures and interactions with other oligos. The probability of non-specific amplification was checked by BLAST (www.ncbi.com/BLAST).

4.2.2. Polymerase Chain Reaction (PCR)

Each DNA sample was amplified in a 25µl of PCR reaction (Table 6) using different annealing temperatures according to the primers (please refer to M&M: 4.2.1). The amplified fragments were loaded on Agarose gel in Tris Acetate EDTA buffer (TAE) buffer stained with Ethidium

bromide (EtBr). In case of sequencing or cloning, the fragments were cut from gel and purified with Promega Wizard® SV Gel and PCR Clean-Up kit following to the manufacturer's guide. DNA sequencing service was provided by Scientific Park of Madrid (Parque Científico de Madrid).

Table 6. PCR components

Component	1 Reaction (25µl)	Final Concentration
10X Biotools Colorless buffer	2.5 µl	1X
MgCl2 (50mM)	1.5µl	3mM
dNTP mix (10mM)	0.5µl	0.2mM
Primer Sense (diluted 1/10, 10µM)	1.25µl	1 µM (0.05–1 µM)
Primer Antisense (diluted 1/10, 5µM)	1.25µl	1 µM (0.05–1 µM)
Go TaqDNApol (5u/µl)	0.25	1.25 U
H2O	Up to 25	
DNA (100ng/µl)	1-5 µl	

4.3. Southern blot

This method was used to study the presence of specific viral DNA sequences in genomic samples. The experiment was performed following the procedures previously described (Gil-Ranedo *et al.*, 2015; Ramírez *et al.*, 1996). Briefly, the restricted DNA samples were resolved in 1%-agarose gels, alkali-blotted to nylon membranes (GeneScreen) and filters were hybridized at high stringency with a AAV2 (Hamilton *et al.*, 2004) or MVM full-length probes ^{32}P-labeled by random priming with α-^{32}P- dCTP (Perkinelmer).

4.4. Western blot

Infected cells were washed with PBS and homogenized in Laemmli buffer (1.2x Concentration). Extracts were incubated at 100°C for 10 min. Protein resolution was performed in denaturing conditions (SDS-PAGE) electrophoresis by 8% Acrylamide:Bisacrylamide gels. The resolved proteins were transferred to Nitrocellulose membrane (0.45µm pore) in humid conditions at constant voltage (100V) for 55min. Detection of proteins in the transferred membrane was made according to standard protocol using Tris-Buffered Saline containing 0.1% Tween20 (TBS-T) for washing

steps, TBS-T + 10% FCS for blocking step and TBS-T + 1% FCS + 1% NP40 for antibody incubations. Table 7 shows the concentration of antibodies that were used in this experiment (see table). Secondary antibodies, α-Rabbit-HRP (horseradish peroxidase) (Jackson ImmunoResearch) and α-Mouse-POD (peroxidase) (Jackson ImmunoResearch), were used in a constant concentration of 1/10000. ThermoFisher Scientific product "SuperSignal West Pico PLUS Substrate" was used to develop the cheminiluminiscence assay to be exposed on X-ray films. Finally, the X-ray films were scanned using Densitometer GS900 BioRad and lane quantification analyses were done using the "Image Lab BioRad" Software.

Table 7. Antibodies used for western blot

Antibodies	Concentration	Source
Anti-VPs	1/4000	Rabbit
Anti-VPsVBK *	1/2000	Rabbit

*Anti-VPsVBK is a purified version of Anti-VPs

4.5. Analysis of viral particles configuration by agarose-blot

An innovative method was set up in order to detect viral particles in their native/denatured state by agarose electrophoresis and blotting to nitrocellulose membranes.

4.5.1. Resolution of viral particles by agarose electrophoresis and blots

Virus samples were prepared in PBS supplemented with 15% glycerol on ice. In case of cell extracts, the extracts were supplied with protease inhibitors (0.1mM phenylmethylsulfonyl fluoride, 10µg/ml Leupeptin, 10µg/ml Pepstatin, 10µg/ml aprotinin *and* 10µg/ml TPCK) and 0.1% Triton in order to remove any innate cellular protease activity (Gil-Ranedo *et al.*, 2015). Samples were loaded in precooled 1.2 % agarose gel at 4 °C prepared in Tris Acetate EDTA (TAE) buffer. The electrophoresis was performed at 4 °C in the same buffer running for two hours at 80V. Next a capillary-blot was performed at room temperature to transfer the virus samples to nitrocellulose membranes. For this, a saline-sodium citrate SSC-10X (1.5M NaCl, 0.15M trisodium citrate) and 10 mM Tris pH 7.5 buffer were used. The agarose gel was placed upon Watman filter, subsequently the nitrocellulose membrane was placed on the top of the gel with two other Watman filters upon it, and finally a stack of filter papers was placed on top of all. Blot was allowed to proceed overnight, and then the nitrocellulose filter was placed on filter paper wet with Tris Buffer Saline (TBS) for 10 minutes at room temperature.

4.5.2. Analysis of native viral particles resolved by agarose blots.

Filters were soaked in blocking buffer containing TBS (50mM Tris-HCl pH:7.5 and 150mM NaCl), 10% FCS and 0.1% tween for 1- 2 hours. Then three steps of washing with TBS-T (TBS+ 0.1% tween) were performed, and the membrane was incubated with the primary antibody in TBS-T supplemented with 1% FCS and 1% NP-40 at 4 °C overnight. Afterwards another three steps of washing at 4 °C were performed and a Horseradish Peroxidase (HRP) secondary antibody (diluted in the same buffer as the primary antibody) was added for 1 hour. Finally, the membrane was washed three times in TBS+ 0.1% tween and once in TBS for 15 minutes and developed by enhanced chemiluminescence (ECL) buffer (Amersham) for 5 minutes. The luminiscence signal was revealed on autoradiography films after several times of exposure.

4.5.3. Analysis of capsid subunits of viral particles by agarose-blots

This method was set up to detect the unassembled capsid subunits irrespective of viral particle configuration. After the overnight blot, the membrane soaked in TBS was placed at 78 °C in an oven for 2 hours. Then it was placed in humid chamber onto a filter paper that contains the solution of TBS + SDS 1%, at 37 °C for 15-20 minutes. These treatments were on purpose to fully denature viral capsids and their protein subunits. Afterwards, the same protocol as native blot was applied to develop the membrane with antibodies, but normally the primary antibody was the anti-VPs (polyclonal) recognizing mainly denatured capsid proteins. Table 8, outlines the antibodies used for native or denatured structures of viral proteins resolved by agarose blots.

Table 8. Antibodies and dilutions used for the agarose blot method

Name	Type	Description	Source	Dilution
Anti VPs	Primary	Denatured condition	Rabbit	1/5000
Anti B7	Primary	Native condition	Mouse	1/2000
Anti-Capsid	Primary	Native condition	Rabbit	1/4000
Anti-Mouse HRP	Secondary	HRP(Thermofisher)	Goat	1/10000
Anti-Rabbit HRP	Secondary	HRP(Thermofisher)	Goat	1/10000

5. Bioinformatic studies

Comprehensive search for endogenous parvoviral elements was initially performed using eight representative members of *Parvoviridae* (MVM, AAV2, BPV, PBoV4-1, B19, PARV-4, AMDV-G and GmDNV) against a total of thirty two genomes belonging to human and non-human primates (Table 9), which are currently available in refseq_genomics and whole genome shotgun (wgs) databases. After obtaining the initial molecular results by representative members of *Parvoviridae,* the seach was extended to all the members of the family (Table 10). Different softwares were used to identify the corresponding DNA sequences of *Parvoviridae* in *Primates* as explained below.

Table 9. *Haplorhini* Suborder list of species tested

Common name	Species	Genus	Family	BioSample	NCBI IDs
Ma's Night Monkey	Aotus nancymaae*	Aotus	Cebidae		
Common Marmoset	Callithrix jacchus*	Callithix	Cebidae	Blood	
Black-capped Squirrel Monkey	Saimiri boliviensis boliviensis*	Saimiri	Cebidae		
Black-capped Capuchin Monkey	Sapajus apella	Sapajus	Cebidae	Blood	
Sooty Mangabey	Cercocebus atys*	Cercocebus	Cercopithecidae	Liver	NW_012004166.1 NW_012006154.1
Golden-bellied Mangabey	Cercocebus chrysogaster	Cercocebus	Cercopithecidae	Blood	
Tanzania Syke's Monkey	Cercopithecus mitis monoides	Cercopithecus	Cercopithecidae	Blood	
Diana Monkey	Cercopithecus diana	Cercopithecus	Cercopithecidae	Liver	
Green Monkey	Chlorocebus sabaeus*	Chlorocebus	Cercopithecidae	Cell lines	NW_006735234.1
Tanzania Black-and-white Colobus Monkey	Colobus angolensis palliatus*	Colobus	Cercopithecidae		NW_012115044.1
King Colobus	Colobus polykomos	Colobus	Cercopithecidae	Blood	
Black Crested Mangabey	Lophocebus aterrimus	Lophocebus	Cercopithecidae	Blood	
Long Tailed Macaque	Macaca fascicularis*	Macaca	Cercopithecidae	Cerebellum Blood	
Northern pig-tailed macaque	Macaca leonina	Macaca	Cercopithecidae	Blood	
Rhesus monkey	Macaca mulatta*	Macaca	Cercopithecidae		
Pig-tailed Macaque	Macaca nemestrina*	Macaca	Cercopithecidae		
Drill	Mandrillus leucophaeus*	Mandrillus	Cercopithecidae		NW_012100098.1
Mandrill	Mandrillus Sphinx	Mandrillus	Cercopithecidae	Blood	
Proboscis Monkey	Nasalis larvatus*	Nasalis	Cercopithecidae		
Olive Baboon	Papio anubis*	Papio	Cercopithecidae	Cerebellum Liver	NW_003872382.1 NW_003876533.1
Hamadryas Baboon	Papio hamadryas	Papio	Cercopithecidae	Blood	
Red-shanked Douc Langur	Pygathrux nemaeus	Pygathrix	Cercopithecidae	Blood	
Golden Snub-nosed Monkey	Rhinopithecus roxellana*	Rhinopithecus	Cercopithecidae		
Gelada Baboon	Theropithecus gelada	Theropithecus	Cercopithecidae	Blood Liver	
Phayre's Leaf Monkey	Trachypithecus phayrei	Trachypithecus	Cercopithecidae	Blood	
Lowland Gorilla	Gorilla gorilla gorilla*	Gorilla	Hominidae	Blood	
Human	Homo sapiens*	Homo	Hominidae	Blood	
Bonobo	Pan paniscus*	Pan	Hominidae		
Central Chimpanzee	Pan troglodytes*	Pan	Hominidae	Blood	
Sumatran Orangutan	Pongo abelii*	Pongo	Hominidae	Blood	
Northern White-cheeked Gibbon	Nomascus leucogenys*	Nomascus	Hylobatidae		
Philippine Tarsier	Tarsius syrichta*	Tarsius	Tarsiidae		

Table 10. ssDNA viral references used for the screening of EVE in primates genomes

Genus	Species	Strain	Acronym	Accession #
Ambidensovirus	*Lepidopteran ambidensovirus 1*	Galleria mellonella densovirus	GmDV	L32896
Amdoparvovirus	*Carnivore amdoparvovirus 1*	Aleutian mink disease virus	AMDV	M20036
Bocaparvovirus	*Ungulate bocaparvovirus 1*	Bovine parvovirus	BPV	M14363
	Ungulate bocaparvovirus 5	Porcine bocavirus 4-1	PBoV4-1	JF429835
Dependoparvovirus	*Adeno-associated dependoparvovirus A*	Adeno-associated virus - 2	AAV2	AF043303
		Adeno-associated virus - 7	AAV7	AF513851
		Adeno-associated virus - 8	AAV8	AF513852
	Adeno-associated dependoparvovirus B	Adeno-associated virus - 5	AAV5	AF085716
	Anseriform dependoparvovirus 1	Goose parvovirus	GPV	U25749
	Anseriform dependoparvovirus 1	Muscovy duck parvovirus	MDPV	U22967
	Avian dependoparvovirus 1	Avian adeno-associated virus strain DA-1	AAAV	AY629583
	Beared Dragon dependoparvovirus	Bearded dragon parvovirus	BDPV	KP733794
	Caprine Adeno-associated virus 1	Adeno-associated virus-Go.1	AAV-Go.1	DQ335246
	Chiropteran dependoparvovirus 1	Bat adeno-associated virus YNM (BtAAV)	BtAAV	GU226971
	Squamate dependoparvovirus	Snake adeno-associated virus	SAAV	AY349010
Erythroparvovirus	*Primate erythroparvovirus 1*	Human parvovirus B19 isolate J35	B19V-J35	AY386330
Protoparvovirus	*Rodent protoparvovirus 1*	Minute virus of mice	MVM	J02275
Tetraparvovirus	*Primate tetraparvovirus 1*	Human parvovirus 4 G1	PARV4G1	AY622943

5.1. NCBI

The National Center for Biotechnology Information (NCBI) website (https://www.ncbi.nlm.nih.gov/) was used for various studies as follows:

5.1.1. GenBank

This database (https://www.ncbi.nlm.nih.gov/genbank/) was used to find the complete genome sequences of annotated Parvoviruses as well as primates.

5.1.2. BLAST

The Basic Local Alignment Search Tool (BLAST) application is an online alignment software (https://blast.ncbi.nlm.nih.gov/Blast.cgi) which was used for comparison of nucleotide sequence against other nucleotide sequence (BLASTn), protein sequence against nucleotide sequence (tBLASTn), protein sequence against protein (BLASTp) and nucleotide sequence against protein sequence (BLASTx) (Altschul *et al.*, 1990; Belshaw and Katzourakis, 2005; States and Gish, 1994). It was also used to check the specificity of designed primers (refer to M&M: 4.2.1). In case of screening, all searches maintained default parameters being suggested by the program. In order to evaluate the significance level of results, a threshold of <1e-10 was set for the E-value associated to the alignment obtained by subject (host) sequence.

5.2. Multiple sequence alignments and phylogenetic analysis

The identified SDEs were aligned along with representative members of correspoding parvovirus sequence using ClustalW (version 2.1) and Jalview (version 2.8.2) programs (Clamp *et al.*, 2004; Sievers and Higgins, 2014). The maximum likelihood phylogenetic tree (Saitou and Nei, 1987) was generated by the MEGA version 7.0.14 software package (Tamura *et al.*, 2013) using the Tanimura-Nei model with 1,000 bootstraps, in order to assess the robustness of the interior branches of the tree (Guindon and Gascuel, 2003).

5.3. Gene Runner

This offline software (Gene Runner version 5.0. 63 beta) was used to design primers and study the interaction of oligos. It was also used to translate the nucleotides to amino acids and find the adequate restriction enzyme to cut the desired sites of a genome.

5.4. RepeatMasker

This online program (http://www.repeatmasker.org/) was used to identify known repeated sequences inside of an unknown sequence (Tarailo-Graovac and Chen, 2009). It was used to distinguish the flanking repeated sequences that surround the endogenous parvovirus sequence inside the primate genome.

5.5. European Nucleotide Archive (ENA)

Sequences obtained by this study were submitted to this database belonging to EMBL (European Molecular Biology Laboratory) by the following link: https://www.ebi.ac.uk/ena

Results

Results

Section A: Post-translational modifications of Parvovirus MVM capsid and cell entry.

1. Virus Titration and specific infectivity of MVMi and MVMt.

Various conventional methods were used to titrate virus as physical particles and infectious units (refer to M&M: 2.3). In these determinations different viral units per milliliters were obtained such as Plaque Forming Unit (PFU), Hemagglutination Unit (HAU), and genomic DNA copy number (DNA). Moreover, the number of particles per each field at 12K magnitude obtained by transmission electron microscope (12k-TEM) was used to study physical differences between viruses. Various stocks of viruses were used along this study from independent infections of NB324K host cells by MVMi and MVMt viruses, and the virus yields were purified by density gradients as described (refer to M&M 2.2). The obtained purified stocks showed different concentration of viral particles and infectious units (Table 11). A ratio was applied between the different methods of titration for each stock.

Table 11. Purified stocks of MVMi and MVMt viruses used in this study.

Stock Label	Extraction date	PFU/ml	HAU/ml	DNA/ml	12k-TEM
MVMi-I	July/2014	1.63E+07	1.70E+04	3.45E+09	
MVMi-II	Sep/2014	2.18E+09	1.02E+06	1.25E+11	1.41E+03
MVMi-III	Nov/2015	6.00E+08	1.25E+05		2.34E+02
MVMi-IV	May/2017	1.50E+09			
MVMi-V	Jul/2017	1.83E+07	1.28E+04		
MVMt-I	Apr/2012	1.00E+11	5.12E+05	2.84E+10	8.72E+02
MVMt-II	Dec/2014	1.51E+09	2.65E+03	4.47E+08	
MVMt-III	Jul/2017	1.07E+09	6.40E+03		

Titers are indicated by the different methods used: Plaque assay (PFU), Hemagglutination (HAU), Real time PCR (DNA), Transmission electron microscopy at 12K magnitude (12K-TEM) *MVMi-III has undergone two rounds of CsCl equilibrium centrifugation. MVMi-V is MVMi-IV subjected to an additional round of centrifugation in CsCl and MVMt-II and MVMt-III are rebanded from MVMt-I.

2. Specific infectivity of virus stocks.

The specific infectivity was measured by dividing the infectious determinations (PFU) versus the methods determining physical particles (HAU, DNA and 12k-TEM). Figure 9 contains column graphs showing the specific infectivity that was obtained by each method. These studies of specific infectivity were performed with all stocks and several times per method. As shown in the figure, the MVMt revealed higher ability to infect than MVMi (almost 2 Log_{10}) in all the studies. Thus, specific infectivity of MVMt was between 55 to 316 times higher than MVMi, which means that less particles of MVMt are needed to form a lysis plaque. These data demonstrated the high efficiency of MVMt infection for further implications.

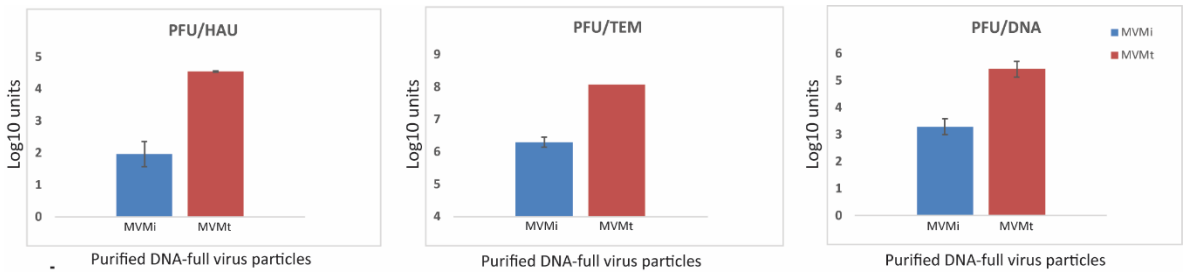


Figure 9. <u>Specific infectivity of viral stocks.</u>
The results were measured by different methods and are shown as Log10. As illustrated in the figure the MVMt shows almost 2 Logs more specific infectivity per particle than MVMi.

3. Analysis of MVMi and MVMt responses to heat.

To search for physical differences between the MVMi and MVMt, highly purified virus particles and capsids were subjected to heat treatments. Their integrity, configuration, and genome uncoating were determined by different methods mentioned below.

3.1. Study of MVM physical particles in response to heat by TEM

The thermostability of the viruses was visualized by TEM (refer to M&M: 2.3.3). The number of particles was estimated and their structural integrity was inspected.

3.1.1. Visualization of the Integrity of the MVM particles

In order to study the integrity of viral particles, highly purified empty capsids and DNA-filled viruses of MVMi and MVMt were subjected to negative staining with Uranyl Acetate. Figure 10

shows 100k magnitude of preparations obtained by TEM. It can be clearly observed that in MVMi empty capsid samples the stain penetrated inside the particles (positive staining), while both wild type and mutant DNA-filled virus particles remained unstained (designated as negative staining).

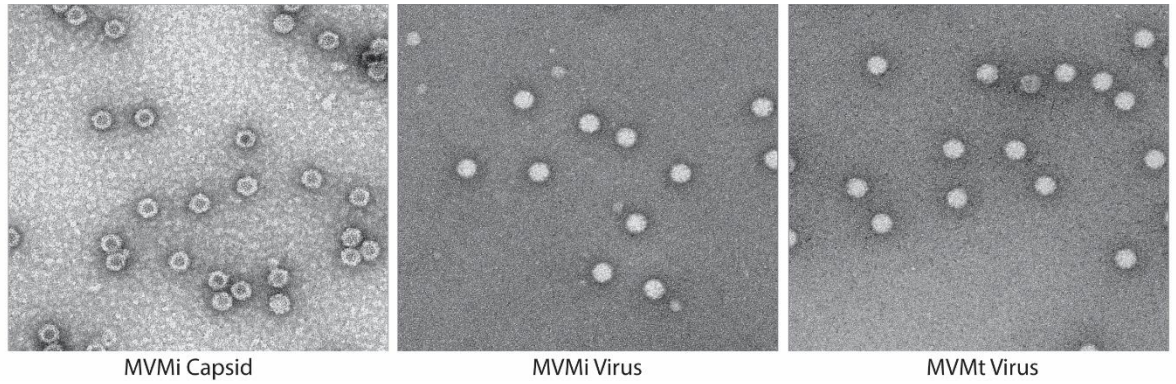


Figure 10. Visualization of MVM particles by TEM.
MVMi capsid, MVMi full virus and MVMt full virus were applied on grids and next subjected to negative staining by Uranyl Acetate. Grids were inspected by JEOL (JEM1010) and images were taken at 100k magnitude. The presence of stain inside the particles is evident only in empty capsids.

3.1.2. Analysis of MVM viruses thermostability

The first analysis seeking biophysical differences between the MVMi and the MVMt was addressed measuring their integrity by TEM in response to heat. For this, equal amounts (HAU) of MVMi and MVMt purified virions were incubated in water bath, for 10 minutes, at temperatures 55°C, 65°C, 70°C, 75°C, 80°C, and 85°C, then immediately placed on ice. One sample of each virus was kept on ice (4 °C) as control. Samples were applied on grids and stained within one hour post-treatment. Under TEM, the number of viral particles were counted per each field at 30k magnitude, using two grids per sample. The number and the staining of particles were captured at different temperatures and summarized in Figure 11.

The electron micrographs of each virus (Figure 11A) illustrate the increase in the proportion of stained particles as the temperature raised. This correlation indicates that empty capsids, or damaged particles permissive to inside staining are generated by heat. Although the original stocks of MVMt and MVMi had a similar low proportion of stained particles when handled at 4 °C (presumably contaminant of empty capsids), the proportion of stained particles is higher for MVMi than for the MVMt as the temperature was raised. For instance, at 80 °C there is no MVMt, but

MVMi harbors countable particles that are mostly stained (65% of all particles) suggesting that the structure of MVMi has more thermostability although particles are emptied or damaged.

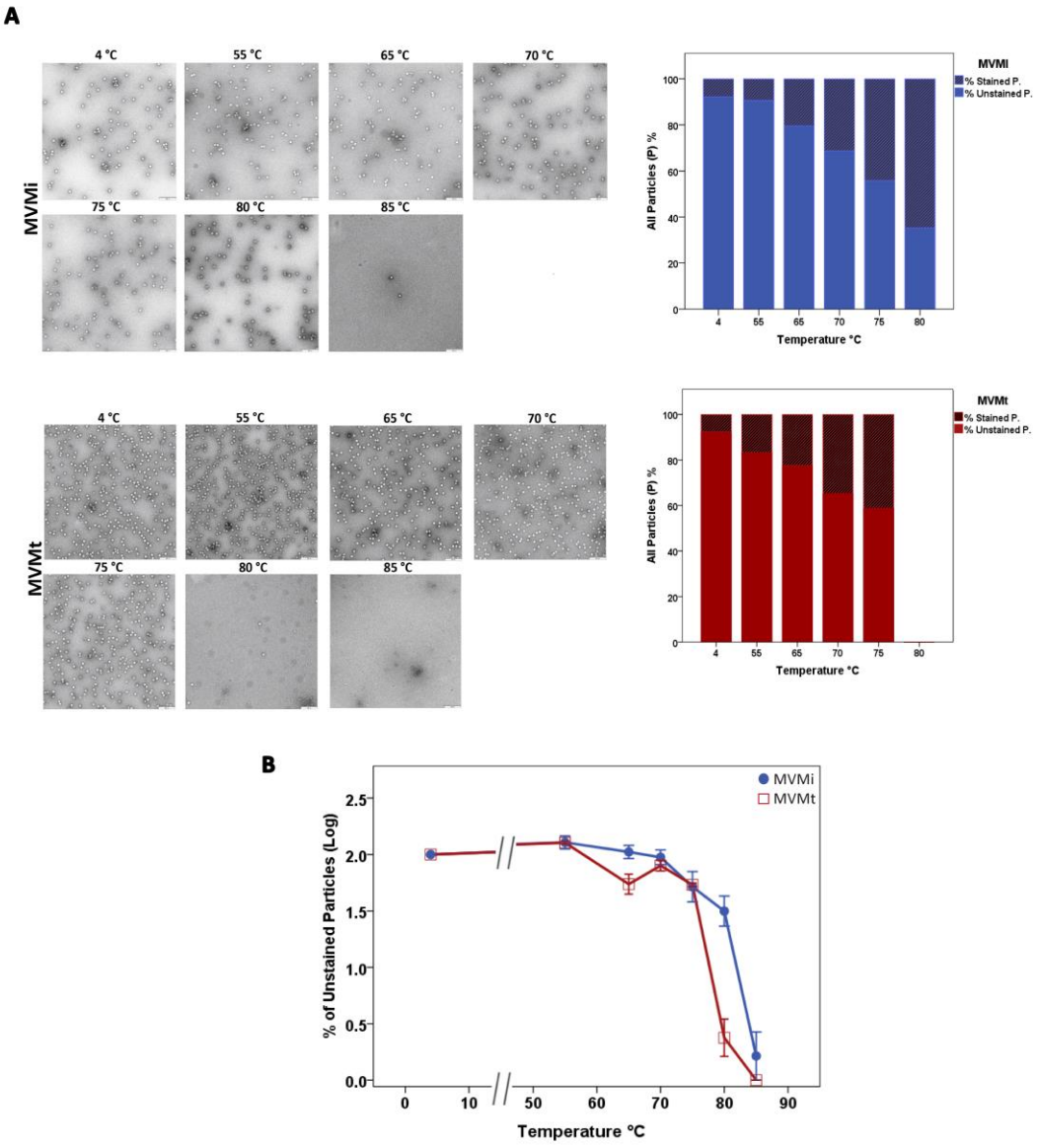


Figure 11. <u>TEM analysis of the MVMi and MVMt thermostability.</u>

A: Electron micrographs (30K magnitude, obtained by TEM) of heat-treated viruses obtained at the indicated temperatures. The graphs to the right show the ratio of stained (presumed emptied or damaged viral particles) versus total viral particles at each temperature obtained in the micrographs. **B:** Graph showing the percentage of unstained particles (presumably intact virions) at each temperature.

Figure 11.B shows a representation of the unstained MVMi and MVMt physical particles (intact virus) behavior during heat treatment. The number of intact viral particles remained similar across low temperatures up to 75 °C, but at 80 °C a drastic difference was noticed between the two viruses: at this temperature a high proportion of MVMi viral particles remained countable, whereas most MVMt virions were not apparent as physical particles. At the higher temperature (85°C) both virions were essentially undetectable. This first analysis showed a clear difference in the physical resistance of the MVMi and MVMt viral particles to heat. The tolerance to stain penetration inside the virions as well as the preservation of the particles integrity drastically differed between the viruses.

3.2. Hemagglutination analysis of virus capsid configuration in response to heat.

The capacity of some viruses to agglutinate red cells (hemagglutination, HA) is mediated by a regular interaction of surface capsid domains with the erythrocyte membrane. This property, which is normally used to determine the amounts of parvovirus particles in viral stocks or virus like particles (VLPs) preparations (Hernando *et al.*, 2000; Jin *et al.*, 2016), was used here to study the changes in the capsid domain(s) triggering HA of the MVMi and MVMt in response to heat. Although the precise amino acid sequence conforming the HA domain has not been determined for MVM capsid, this information is available for a small domain conformed at the surface of the highly related CPV capsid (Chang *et al.*, 1992), suggesting that heated MVM capsids failing to HA may have altered the configuration of some surface domains.

Identical volumes and concentration (HA units) of highly purified MVMi and MVMt DNA-filled virus and empty native capsids were heated in PBS for 10 min at temperatures indicated in Figure 12, cold down immediately on ice, and subjected to a conventional HA assay with mouse erythrocytes. As shown in the figure, both viruses were essentially HA competent at moderate temperatures up to 65 °C. At 70 °C the HA activity decayed similarly in both viruses and for both types of particles, suggesting that the packaged viral genome does not contribute to stabilize the capsid in this assay. At 75 °C a remarkable fraction of MVMi empty and full viral particles remained HA competent, whereas the MVMt particles lost all HA capacity. No HA was detectable at 80 °C for any of the four types of particles assayed. Therefore, the configuration of the HA domains on the MVMi capsid tolerates heat at higher temperature than the MVMt domain(s) does, a property not detected by EM visual inspection.

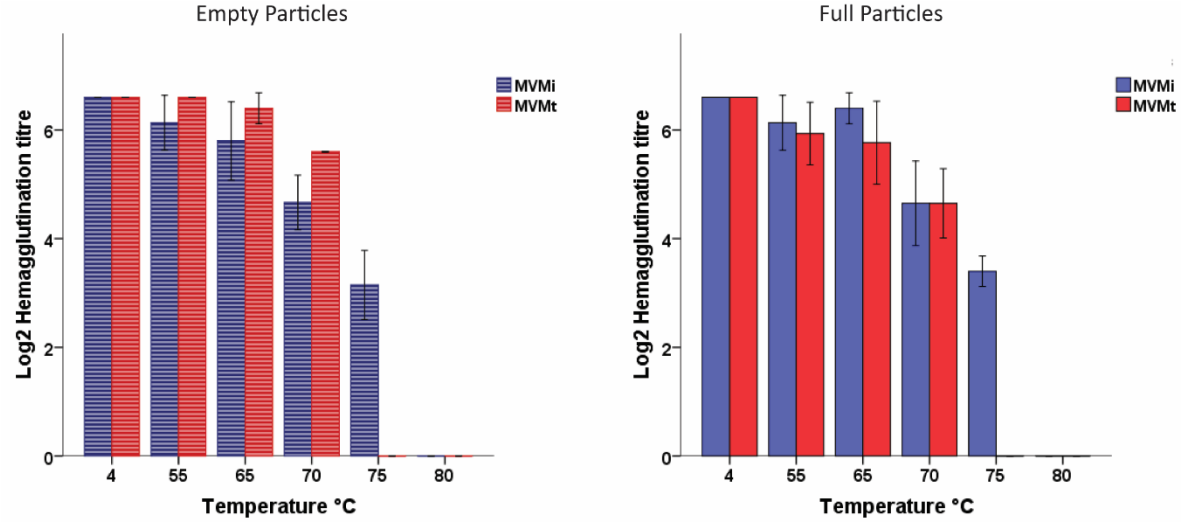


Figure 12. **Heat effect on MVM capsid and virus HA capacity.**
Highly purified full virus and empty capsids of MVMi and MVMt were diluted to 0.5ml of PBS at 4°C. Aliquots of 75µl of each sample were subjected to the indicated temperatures (4°C, 55°C, 65°C, 70°C, 75°C and 80°C) for 10 minutes. A total of 50 µl of each aliquot was used for serial dilution in a conventional hemagglutination assay with 1% mouse erythrocytes overnight. The assay was performed in triplicate, and the remaining HA activity was recorded.

3.3. Heat effect on MVM particles assessed by their behavior in agarose electrophoresis

In order to visualize the integrity and configuration of intact MVM viral particles, a method based on the resolution of viral particles by their electrophoretic mobility in agarose gels and subsequent detection by western-blotting to membranes was developed.

3.3.1. Setting-up the method

The reliability of the method was firstly evaluated using viral particles resolved in 1.2% agarose gels at 4 °C and blotted overnight to nitrocellulose membranes. Blotted particles were then stained with specific antibodies to determine the preservation of capsid configuration and integrity. For this, an antibody raised against denatured MVM capsid proteins (α-VPs), and a monoclonal antibody (α-B7) recognizing an epitope configured in intact capsids, were used under denaturing and native conditions respectively (refer to M&M: 4.5). Figure 13A shows a representative result of this initial test in which MVMi and MVMt empty capsids (C) or DNA-filled viruses (V), blotted to membranes, were developed with these antibodies. A number of conclusions can be drawn:

I) The α-VPs antibody reacts with the structural proteins of denatured as well as intact viral particles.

II) Denatured VPs proteins move faster than intact viral particles, and at a common position in gels regardless their origin (MVMi or MVMt, as well as empty or full virus).

III) Intact capsids of both viruses can be quantitatively detected with the α-B7 antibody, but not if viral particles were previously denatured with SDS and heat. Therefore, the configuration of viral capsids tolerates the agarose-electrophoresis as well as the blot procedure.

IV) In native conditions, the MVMi full virus as well as empty capsid runs faster than corresponding particles of MVMt. This difference is lost when viral particles were heat-disassembled.

V) In the native state of both viruses (PBS, 4°C), full viral particles (V) have a heterogeneous pattern not observed in empty capsids.

VI) Full viral particles (V) run faster in gels than empty capsids (C) do.

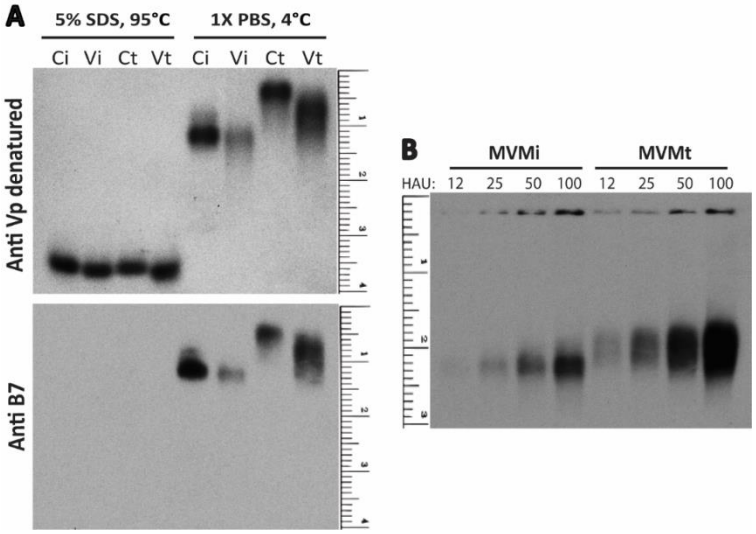


Figure 13. Setting-up the agarose-blot method for MVM particles.
A: Resolution of viral particles under denaturing and non-denaturing conditions. MVMi capsid (Ci), MVMi virus (Vi), MVMt capsid (Ct) and MVMt virus (Vt) were prepared in denaturing (5% SDS heated at 95°C) and non-denaturing (PBS at 4°C) conditions, and subjected to two hours of electrophoresis (80 volts) in 1.2% agarose gel. Samples were capillary-blotted overnight to nitrocellulose and stained with the indicated antibodies. A representative experiment of two is shown. B: Quantitative detection of MVM particles by the agarose blot method. The indicated hemagglutination units (HAU) of purified viral particles of MVMi and MVMt were subjected to the agarose-blot method. Filters were developed with the anti-VPs antibody and a chemiluminiscence method.

3.3.2. Quantitative analysis of capsid stability by the agarose-blot method.

In order to test the reliability of this method in quantitative response, graded amounts of native purified DNA-filled viral particles (determined by an hemagglutination assay) of the MVMi and MVMt were loaded in 1.2% agarose, blotted, and developed under denaturing conditions with the anti-Vps antibody. As observed in Figure 13.B, this method allows a quantitative detection of both types of viral particles. The heterogeneous migration in the samples of the MVMt is particularly evident. The apparent differential detection of the viral particles intensity may reflect the 50% error in estimation of capsid amounts by hemagglutination.

3.3.3. Heat response of viral capsids analyzed by the agarose-blot method.

The differential behavior of the MVMi/MVMt observed by TEM and HA analyses could be accounted by several mechanisms: full capsid disassembly, partial loss of capsid integrity/configuration, or genome ejection. To study the possible mechanisms involved, we analyzed the thermostability of viral capsid configuration by the agarose-blot method. Highly purified MVMi and MVMt viruses were subjected to different temperatures (from 37°C to 85°C) loaded in agarose gels, and blotted to nitrocellulose membranes. Filters were developed with the α-Capsid and α-B7 the antibodies to detect the preservation of the entire capsid configuration or the major 3fold-capsid epitope, respectively. Signal intensity was quantified by ImageJ program. The results are outlined in Figure 14.

Some major comments on the Results obtained by this method are:

-Configuration of capsid epitopes in response to heat.: (i) The overall capsid epitopes of MVMi are more resistant to heat than those of MVMt, as more capsid proteins remained reactive in viral particles to the Capsid-antibody, mainly at 50 °C, but also at higher temperatures; (ii) there is a progressive loss of capsid epitopes in both virus suggesting particles disassembly, withan approximate KD_{50} of 52 °C for MVMt and 62 °C for MVMi; (iii) Consistently, the 3x major epitope recognized by the B7-Mab is much more resistant to heat in MVMi than for the MVMt across the 50-70 °C interval, as judged by the B7-MAb staining of the viral particles. These results strongly suggest that the thermostability of DNA-filled viral capsids is mainly regulated by the configuration of the B7-epitope on the 3x-axis of the capsid, and that this epitope is much more resistant in MVMi than in the MVMt.

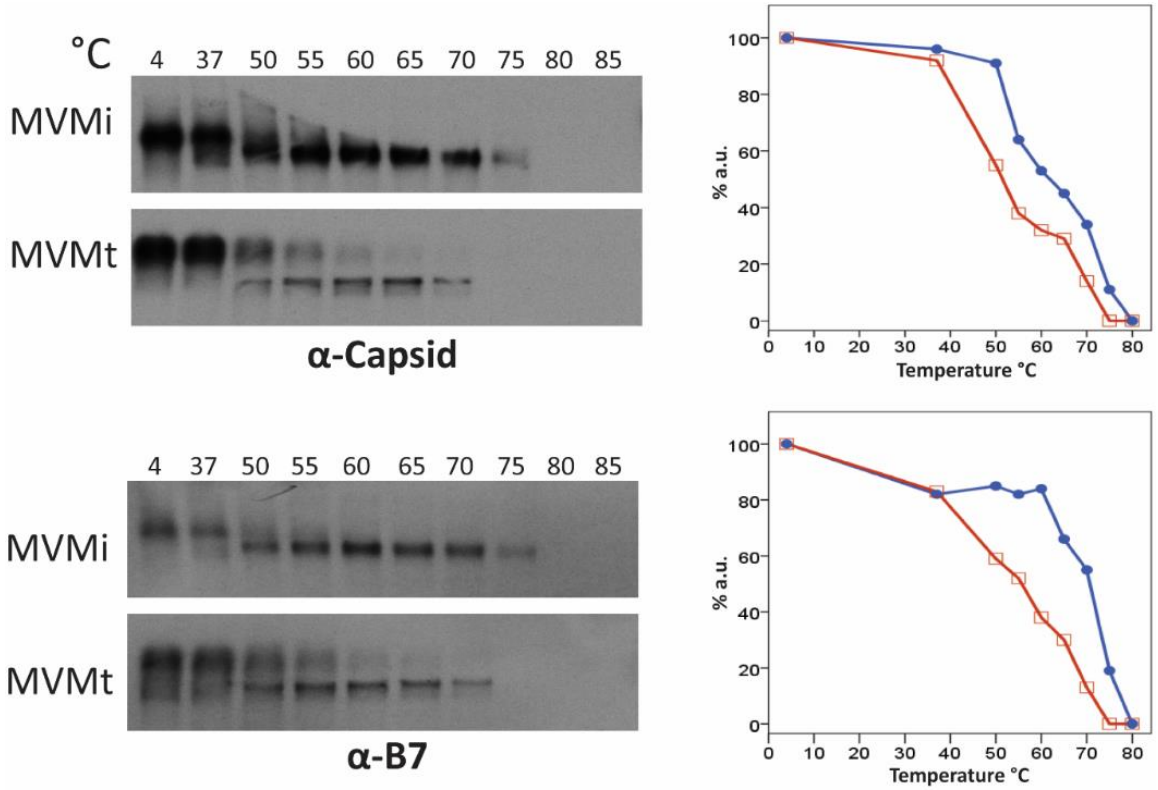


Figure 14. **Agarose-blot analysis of the heat response of viral capsids.**
The figure shows an agarose-blot analysis upon treatment of purified MVMi (blue) and MVMt (red) viral particles to the indicated temperatures for 10 min in PBS. The concentration of viral capsids was normalized by hemagglutination. Filters were proceeded for native staining with the anti-MVM capsid (upper image) or anti-B7 (lower image) antibodies, and the quantitative results are shown to arbitrary units of densitometry in the right. Representative results from two experimental determinations are shown.

- Heat effect on viral particles mobility in gels: (i) There is an irreversible change of mobility in both viruses (manifested as a faster mobility band) happening at 50 °C and maintained at higher temperatures, although this shift may not be completed for the MVMt until 60 °C; (ii) Viral particles in both configurations are equally recognized by the B7-Mab, indicating that the configuration of this 3x-epitope is not related to the mobility shift.

3.4. Virus genome uncoating and infectivity in response to heat

The degree of capsid disassembly of DNA-filled viral particles in response to heat was quantitatively determined by measuring the uncoated ssDNA virus genome susceptible to DNAse. For this, normalized amounts of hemagglutination units of highly purified wt and mutant viral particles were treated 10 minutes under the corresponding temperatures shown in Figure 15.

Samples were next digested with DNase (Promega) for 2 hours at 37°C to remove the exposed DNA, and the remaining protected DNA in the samples was extracted by Qiagen DNeasy Blood and Tissue kit (refer to M&M: 4.1.1). Finally, the viral DNA was amplified by Real Time PCR with specific MVM primers (refer to M&M: 4.2.1). The assay was performed at least three times for each condition to obtain reliable determinations of the amount of exposed viral genomes in each sample with standard deviations.

As shown in Figure 15A, the genome of viral particles maintained at 4 °C are fully resistant to DNase due to natural capsid protection. As viruses were subjected to temperature raise, the genome became progressively exposed to DNase due to loss of capsid protection or uncoating. The curves of uncoating for MVMi and the MVMt showed dissimilar kinetics with major differences between viruses at the 55-70 °C interval. The MVMi kinetics followed a sigmoidal curve suggesting a cooperative process across 50-60 °C, whereas the MVMt kinetics did not adjust to sigmoidal curve, showing increased genome resistance to DNase at this temperature interval, as its genome could be amplified to higher levels than that of MVMi. This result indicated that during viral conformational changes, which lead to uncoating induced by heat *in vitro*, the ssDNA genome is better protected by the MVMt capsid than the MVMi capsid. Therefore, the T/A mutation lowers or restricts the changes in capsid configuration induced by heat that leads to genome uncoating.

To analyze the effects that the physical treatments had on infectivity, which is the major biological virus property, highly purified wt and MVMt DNA-filled viruses were subjected to heat under the conditions described above, and the infectious titer determined by a plaque assay on NB324K cells (see M&M: 2.3.4.1). As shown in Figure 15B, the MVMt maintained infectivity to higher levels than the MVMi across the heat treatment. Interestingly, the MVMi infectivity decayed mainly at the 55-65 °C interval, whereas the MVMt infectivity decay was more progressive and occurred along the entire range of temperatures. Thus, the different response to heat of the virus infectivity fairly reflects the genome uncoating curves showed in Figure 15A.

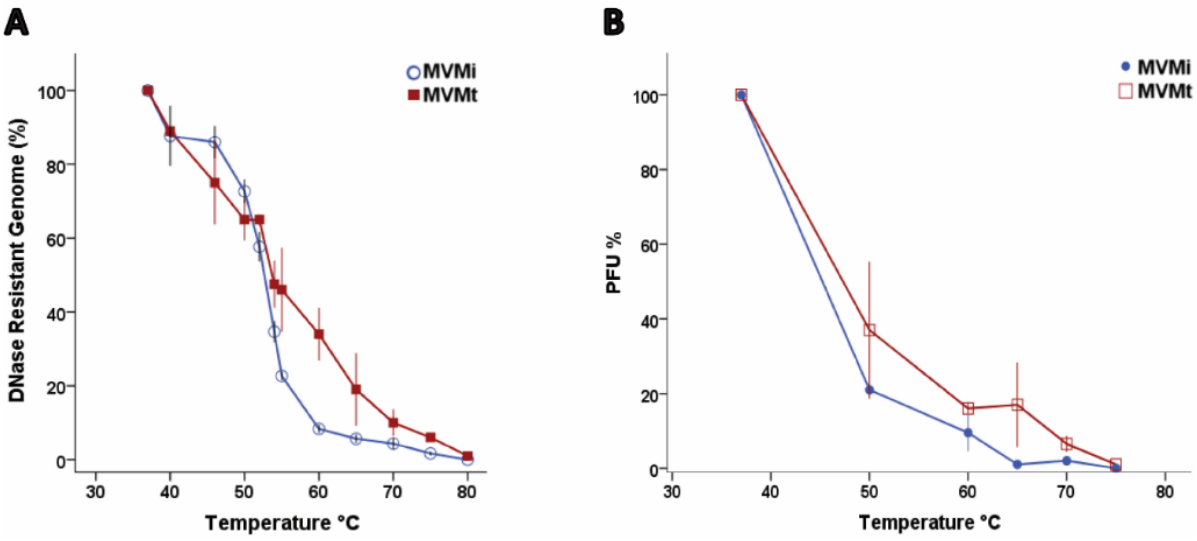


Figure 15. Analysis of MVMi and MVMt genome uncoating and infectivity in response to heat.
A: qPCR quantitative analysis of the DNase-resistant viral genomes upon subjecting purified MVMi and MVMt viral particles to the indicated temperatures for 10 min in PBS. The concentration of viral capsids was normalized by hemagglutination. Average values with standard errors from experiments performed in triplicated with different viral stocks are shown. **B**: Heat effect on MVMi and MVMt infectivity. Viral samples were treated as in A and subjected to a plaque assay on NB324K cells. The results are the mean with standard errors from two independent experiments performed in triplicate.

3.5. Heat effect on VP2/3 cleavage and virus mobility in agarose electrophoresis.

The mature MVM virion exposes about 20 amino acids of the n-terminal sequence (Nt) of the VP2 major structural protein (2Nt) out of the coat, thereby allowing its cleavage by trypsin and other proteases *in vitro* (Sánchez-Martínez *et al.*, 2012 and references therein). We tested whether the 2Nt exposure in response to heat may differentially influence the behavior of the MVMi and MVMT virions in native agarose gels. Highly purified virus particles were heated at serial temperatures for 10 min, then cooled on ice, and subsequently digested with excess of trypsin. Samples were loaded into native agarose gels to test virions configuration (Figure 16A), as well as in SDS-PAGE (Figure 16B) to determine VP1, VP2 and VP3 proteins integrity and resistance to trypsin cleavage.

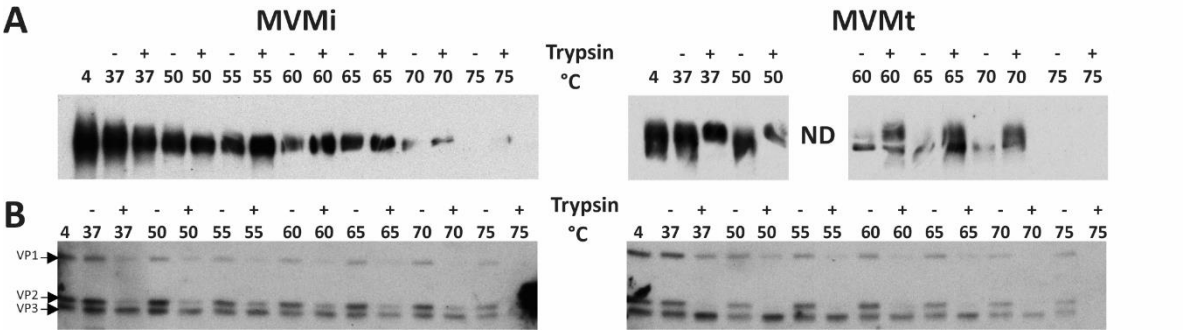


Figure 16. <u>Effect of VP2/VP3 cleavage on MVM mobility in agarose gels.</u>
A: Purified samples of the respective viruses heated at the indicated temperatures were subsequently digested with trypsin at 37 °C and analyzed by the agarose-blot method. Filters were developed under denaturing conditions with the anti-VPs antibody. **B:** Parallel analysis performed with the same samples by conventional SDS-PAGE and western-blot. The results are illustration of two independent experiments.

As shown in Figure 16, both viruses maintain their capsid configurations resolved by agarose up to 70 °C, whereas at 75 °C a complete capsid disassembly allows trypsin digestion of VPs subunits. Also in both viruses, the VP1 became cleaved since 55 °C. In spite of these identities, this analysis allowed to identify some major differences between both viruses that can be summarized as follows:

I) The VP2 protein can be quantitatively cleaved to VP3 in the MVMT virus subjected at any temperature treatment. However, the VP2/3 cleavage in the MVMi virion is more sensitive to heat, indicating a differential exposure of 2Nt in response to heat between both viruses.

II) The trypsin digestion switches the electrophoretic mobility of both viruses in native agarose gels. In particular, trypsin digestion without previous heating (samples remaining at 37 °C) leads to the loss of heterogeneous mobility of virus samples, an effect particularly drastic in the MVMT mutant. This result indicates that the 2Nt exposure is at least partially responsible for the characteristic heterogeneous mobility of the viruses in agarose gels, with more pronounced effect for the MVMT mutant.

III) The MVMT virion is poorly resolved in agarose gels in the range of 50-70 °C, but became focused upon trypsin digestion. This observation suggests that the heated MVMt virion behaves as a particle of heterogeneous configuration, with differentially exposed 1Nt and 2Nt sequences that

affect its navigation through the agarose gel net. This heterogeneity in gel is lost when trypsin cleaves these N-terminal sequences of the VPs subunits.

4. Effect of pH combined with temperature on MVM genome uncoating and infectivity.

Parvovirus entry involves trafficking through the low pH and other stress conditions of the endosomal network. This process is necessary for the viral particles to undergo the structural changes leading to successful infection, but paradoxically it may also inactivate the incoming virus, thereby representing a major barrier against cell invasion. We therefore analyzed the viral particles integrity and infectivity under conditions that may mimic the endosomal environment. For this, highly purified MVMi and MVMt viral particles were subjected to a range of temperatures under neutral (7.5) or a mean endosomal pH value (5.5), and their integrity was assessed by genome uncoating.

The efficiency of uncoating was indirectly studied by DNase digestion of unprotected genomes, and determination of undigested ssDNA was assessed by qPCR with specific MVM primers. Heat may alter the structure of viral capsids so the uncoated DNA can be digested by the DNase. The conditions of the reaction performed to study the thermostability of these viruses (time, pH, DNA extraction method, doses of enzyme, and qPCR) are indicated in the 2.3.2 section of M&M. The Figure 17.A illustrates the results obtained under these conditions. At pH 7.2 the genome uncoating drastically differed between both viruses across the range of temperatures tested. MVMi holds packaged genome efficiently up to 45 °C, but most viral particles drastically uncoat by 55 °C. In contrast, the MVMt genomes remained significantly intact to 60 °C, but lost capsid protection at higher temperatures. When the study was performed at pH 5.5 however, while the kinetic of genome uncoating for the MVMt was similar to that of neutral pH, the MVMi genomes remained essentially full protected by the capsid even at 60 °C. Therefore, a low pH environment, as registered across the endosomal traffic network, remarkably prevents the MVMi genome uncoating *in vitro,* but it does not affect this process in the MVMt.

To seek for a biological relevance of this observation, the same virus samples subjected to pH and temperature treatments were analyzed for infectivity. Figure 17.B shows the plaque forming capacity of the treated viruses. The data shows that the MVMt plaque forming capacity was irrespective of the pH used for the heat treatments. However, the MVMi plaque formation

was highly benefited from a low pH buffering while heating. These results are in perfect correlation with the genome uncoating test performed with these samples outlined in Figure 17.A, indicating that the gained relative infectivity of MVMi at low pH is most likely due to the increased genome protection by the capsid coat.

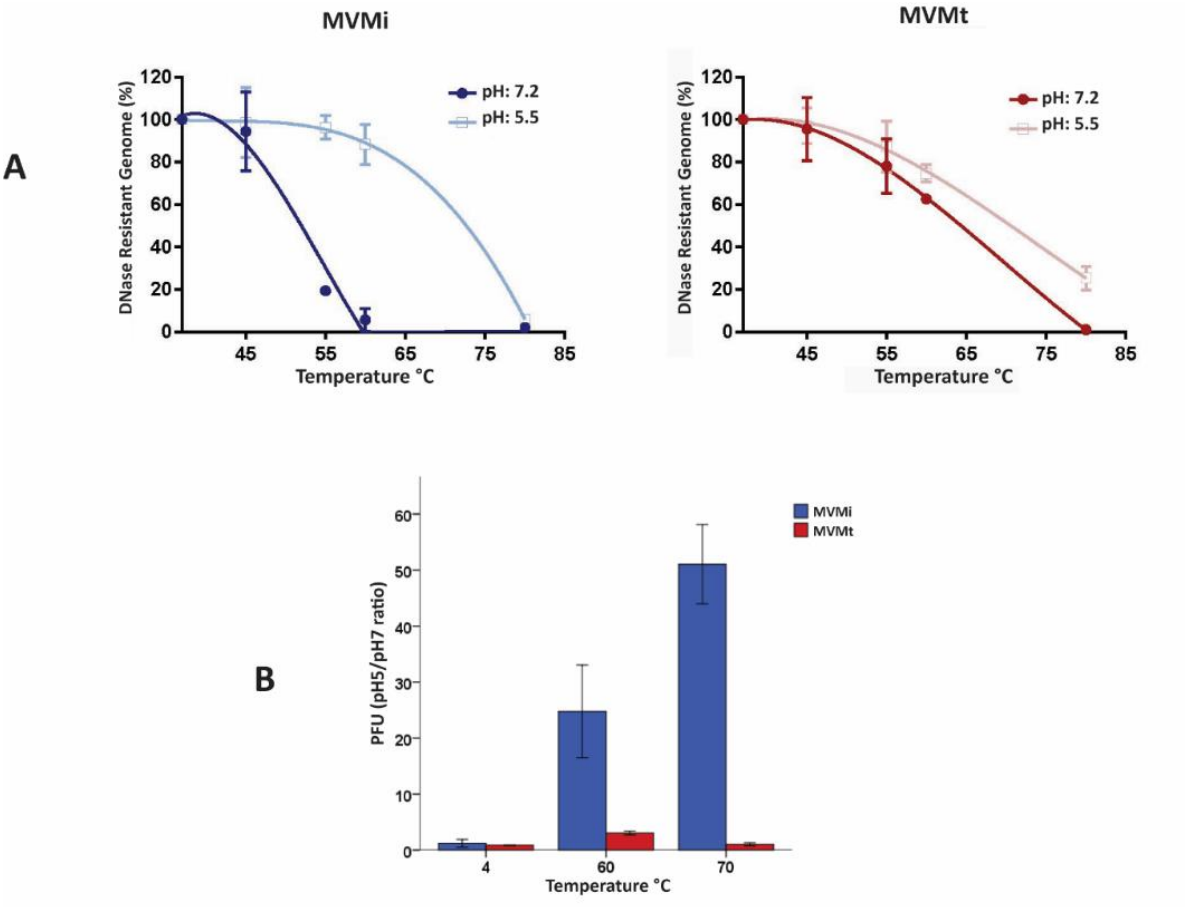
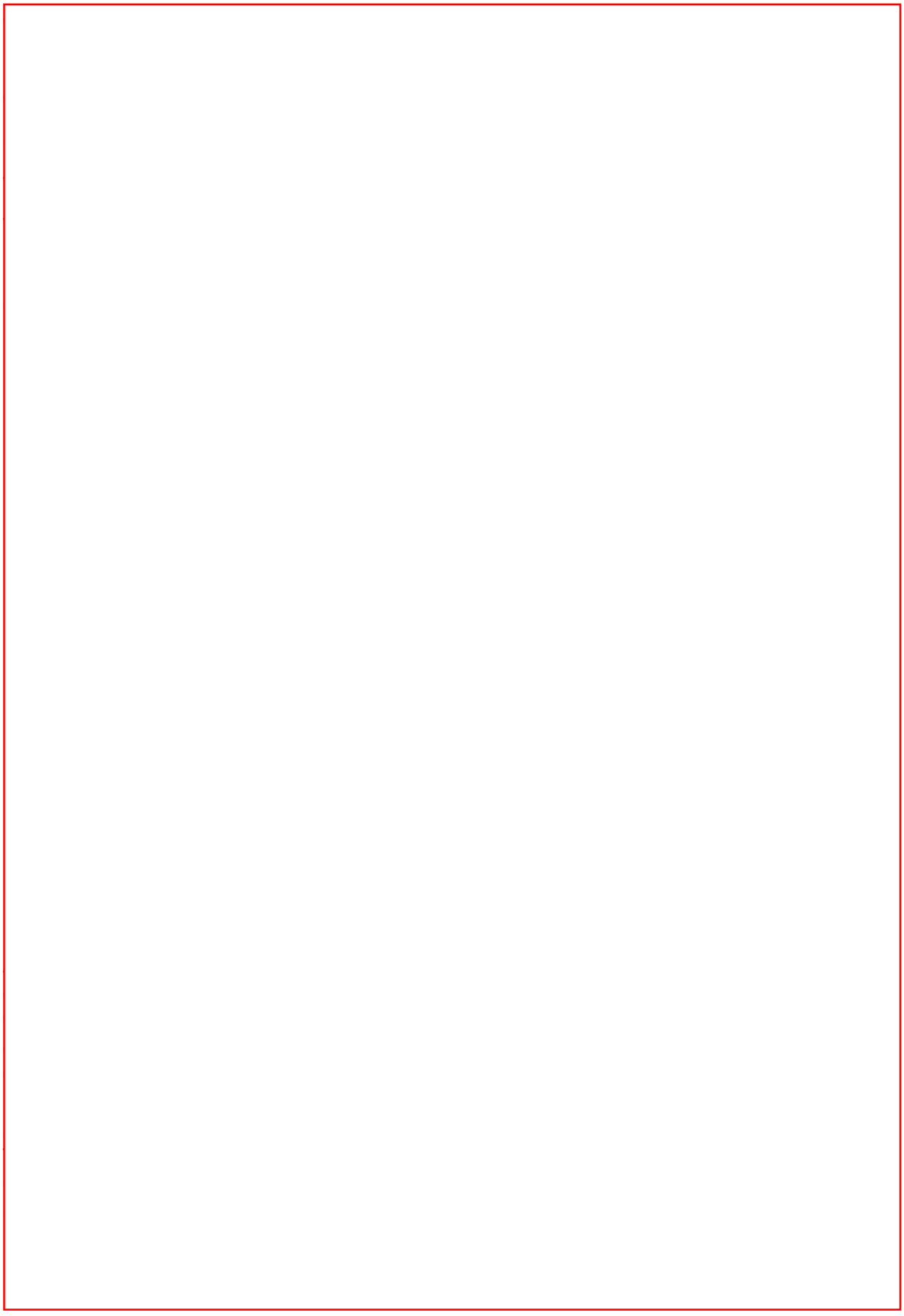


Figure 17. <u>Effect of pH on temperature-induced genome uncoating and infectivity.</u>

A: Purified viral samples were subjected for 10 min at the two pHs and indicated range of temperatures, and subsequently digested with DNase. The remaining DNA was extracted by Qiagen DNeasy Blood & Tissue Kit and analyzed by a qPCR assay. Values are the mean with standard errors from three independent determinations. The thermostability of the MVMi virus in acidic pH is remarkable when compared to that of MVMt. **B:** Effect of temperature under different pHs on MVM infectivity. Purified viral samples subjected to the mentioned pHs and temperatures treatments were inoculated on NB324K cell monolayers, and PFU formation scored at 6 dpi. The ratio of plaque forming capacity (obtained at pH 5.5 versus 7.2 and at the indicated temperatures) is represented. These data are the means with standard errors obtained for three samples per virus from two independent experiments.

5. PROTECTED DATA of the MVMi and MVMt capsids

PROTECTED DATA



6. Role of the endosome in the MVMi and MVMt entry pathway.

The entry process of Parvovirus, particularly in MVM, is very sensitive to the endosomal pH. Indeed, it was described that MVMp requires low pH of the endosome for efficient infection (Ros *et al.*, 2002), and moreover the bulk of the incoming MVMp particles does not escape from the endosome (Garcin and Panté, 2015) (see 4.3. of introduction). However, no data is available for the entry process in the MVMi infection. We therefore tested whether endosomal traffic may underlie the MVMi/MVMt differential specific infectivity.

6.1. Effect of endosomal inhibitors on MVMi and MVMt gene expression.

For this experimental approach we tested successful infections by monitoring the expression of the NS1 protein, the large main cytotoxic product of MVM. NB324K cells were inoculated with purified MVMi and MVMt. The latter virus was used at two multiplicities in order to normalize the infection for the total amounts of viral particles (UHA) or infectious particles (PFU), in respect to the single dose of MVMi that was used. After adsorption, endosomal inhibitors were added to the medium and infection was allowed to proceed for 20 hours. Cells were subsequently inspected for NS1 expression by IF, and collected data were quantitatively recorded in several fields of cells from duplicate experiments. As observed in Figure 19, the percentage of cells expressing NS1 in the presence of 2.5 µM ClQ declined to 50% in case of MVMt, whereas MVMi required a higher ClQ concentration close to 5 µM for a similar inhibition. Likewise, in the presence of 10mM NH_4Cl, the amount of MVMt infected cells expressing NS1 reached 50%, while MVMi required 25 mM. Thus, the MVMt was significantly more inhibited than MVMi by Chloroquine as well as by NH_4Cl. Therefore, both endosomal inhibitors indicated that the MVMt is much more sensitive to the low pH during endosomal sorting for entry than the MVMi virus. These data suggest that capsid post-translational modifications may determine a structural plasticity that is advantageous for endosomal trafficking and successful infection.

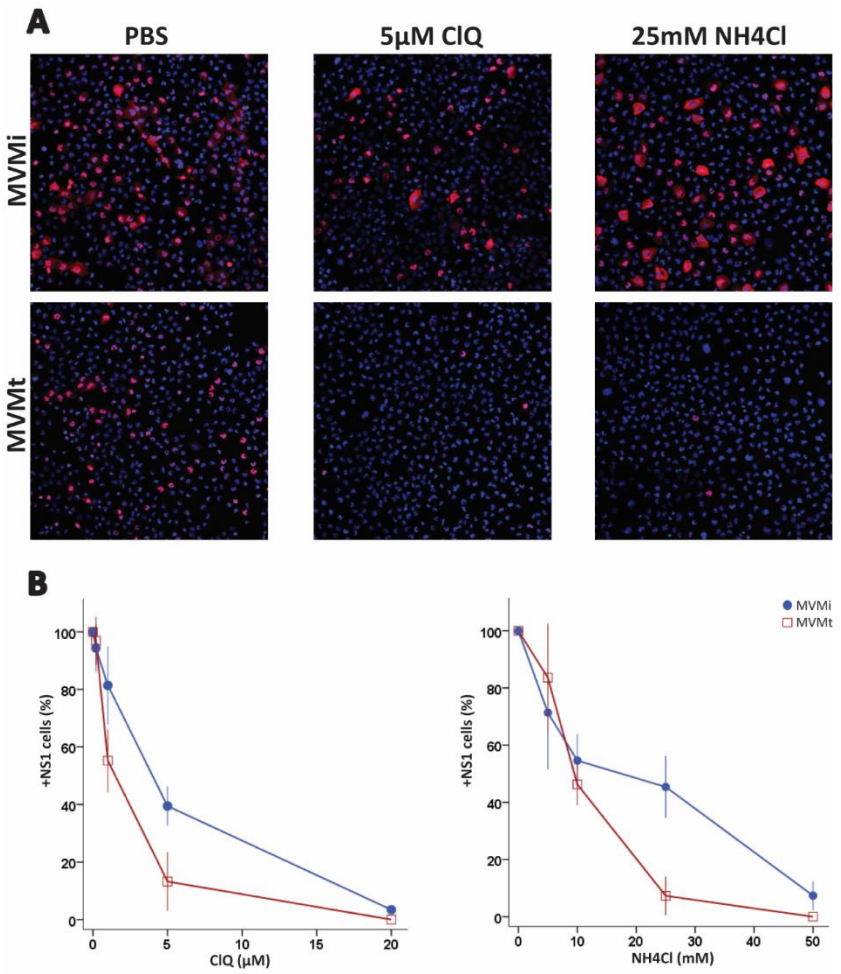


Figure 19. <u>Effect of endosomal inhibitors on viral NS1 protein expression.</u>
Cells seeded onto coverslips were inoculated with normalized amounts of infectious units (PFU/cell) of the MVMi or MVMt, and the NH4Cl and Chloroquine inhibitors were added at the indicated concentration post virus adsorption. **A**: Representative fields of cells stained for NS1 expression at 24 hpi. At least one thousand cells were counted per virus and condition. **B**: Cells undergoing NS1 expression in the presence of endosomal inhibitors. *Right:* Inhibition with NH4Cl. *Left:* Inhibition with Chloroquine doses. Data (average and standard errors) from two independent experiments are shown

6.2. Intracellular features of entering MVMi and MVMt virions.

6.2.1. Capsid subunits and viral genome association of the incoming viral particles.

We sought intracellular features in virions and capsids at early times post infection. NB324K cells were infected with normalized amounts of viruses, and viral samples were harvested at different time points to be resolved under native conditions by the agarose-blot method. Filters were probed (in parallel blots) for VPs and ssDNA viral genome. This study aimed to identify changes in viral capsid integrity/mobility inside the cells and their association to the viral genome. Heat-denatured viral control samples showed released genome (Figure 20, 90°C lower filter), and the

migration of the disassembled capsid proteins as a faint band at the front of the gel (Figure 20, 90°C upper filter). As expected, the MVMi and MVMt heat-uncoated ssDNA genomes migrate at the same position in the gel. However, in native virion samples, the genome migrates at the same position of the corresponding VPs, indicating capsid packaging. As outlined above, the modifications of capsid subunits alter the characteristic mobility of intact viral particles (MVMt virions moving slower than MVMi), and therefore of the packaged genomes.

Viral samples were harvested along hours post-infection (HPI) to be analyzed by this method. As observed in Figure 20, the pattern of capsid proteins-viral genome association from each virus remained at the characteristic mobility in all samples. The resolved viral genomes migrated at the position of capsids during early stages of virus entry. This ssDNA-VPs association suggests that a significant fraction of the genome is maintained packaged within capsids along the entry process., although the extent of this process may vary between viruses (unclear study). At late times of the infection (16 hpi) the synthesis of VPs and of viral DNA is patent in both viruses, denoting their high infectivity in this cell system.

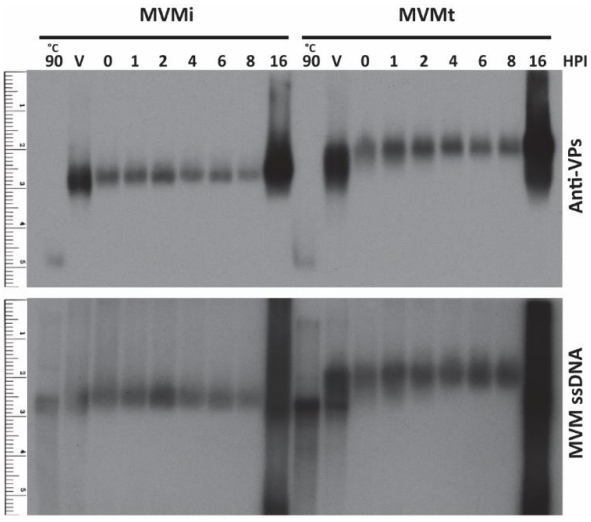


Figure 20. <u>Viral genome association to VPs along early times of MVM infection.</u>
NB324K cell monolayers were inoculated with normalized amounts of purified viruses (180 UHA/10e5 cells). The samples of homogenized cells, collected in PBS at different hours post-infection (hpi), were resolved by native agarose-blot gels. Each sample was simultaneously loaded on two equivalent agarose gels running in parallel at 4 °C. *Upper*, blots of viral samples probed with anti-VPs antibody under denaturing conditions (see MM). *Lower*, agarose gels blotted to nylon membrane by alkali method, showing the ssDNA genomes resolved by hybridization with a ^{32}P-labeled MVM probe (see MM). The co-migration of the VPs and ssDNA in the gels is illustrated by the rulers (left margins). Virus particles in their native configuration at 4 °C (V) and heat disassembled (90 °C) were loaded as control.

6.2.2. Viral particles configuration along the entry pathway.

The co-migration of VPs with genome ssDNA in gels suggested packaging, but did not inform about the configuration of the viral capsid. This issue was addressed by studying changes in the capsid epitopes along early stages of virus cell entry. Purified virus or empty capsids applied on NB324K cells were sampled along the time and loaded in agarose gels under native conditions as above. But blots were probed with the anti-MVM capsid antibody that mainly recognizes capsid configuration epitopes. Figure 21.A shows that MVMi and MVMt empty capsids progressively lost antibody recognition from 2 to 8 h post-adsorption. The same pattern of loss of antibody recognition along the time was observed for the MVMt virus samples Figure 21.B, as well as in MVMi virus samples (data not shown). We therefore conclude that although the ssDNA genome remains likely packaged, the entering viral capsids undergo conformational changes along the endosomal pathway involving the loss of capsid configuration epitopes.

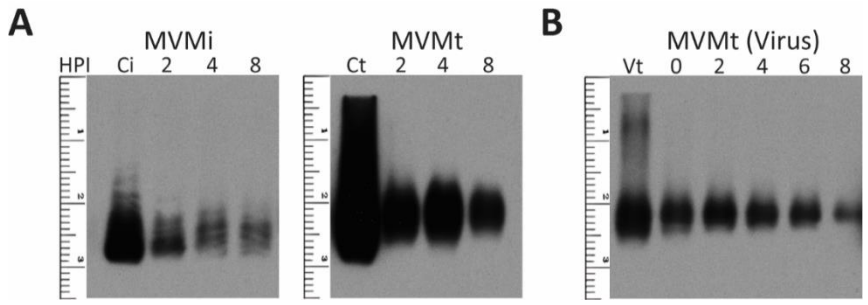


Figure 21. <u>Configuration of capsid epitopes in trafficking virus.</u>
NB324K cell monolayers inoculated and sampled as above were loaded on agarose gels and run at 4 °C. Gels were blotted to nitrocellulose filters under native conditions and probed with the anti-MVM capsid antibody. **A:** Empty capsids sampled at the indicated hours post-adsorption**. B:** Purified viral particles of MVMt analyzed post-adsorption time. Filters were developed (ECL, Amersham) and exposed to autoradiography for 5 min.

6.2.3. Confocal analysis of intracellular virus traffic.

Putative differences between the intracellular traffic of both viruses were finally studied by confocal microscopy. For this, NB342K cells monolayers inoculated with normalized amounts of purified viral particles were fixed at serial time points, and capsid proteins and DNA of the incoming viral particles were stained by IF. Figure 22 shows a panel of representative fields of these experiments. MVMi and MVMt viral particles could be identified at 0 hpi post-adsorption by a faint DNA staining. As the infections proceeded, the staining with all markers (intact Capsid/capsid proteins/viral DNA) significantly increased (Figure 22, panels 2 and 6 hpi), denoting

a presumed accumulation of incoming viral particles in endosomal vesicles. The signals under confocal inspection were consistently higher for the MVMt versus the MVMi, which may suggest a higher intracellular stability of this virus. Capsid proteins and DNA of both viruses were detected by high accumulation at 22 hpi in most cell's nuclei, reflecting a *de novo* synthesis of viral macromolecules by the high multiplicity of infectious virus used.

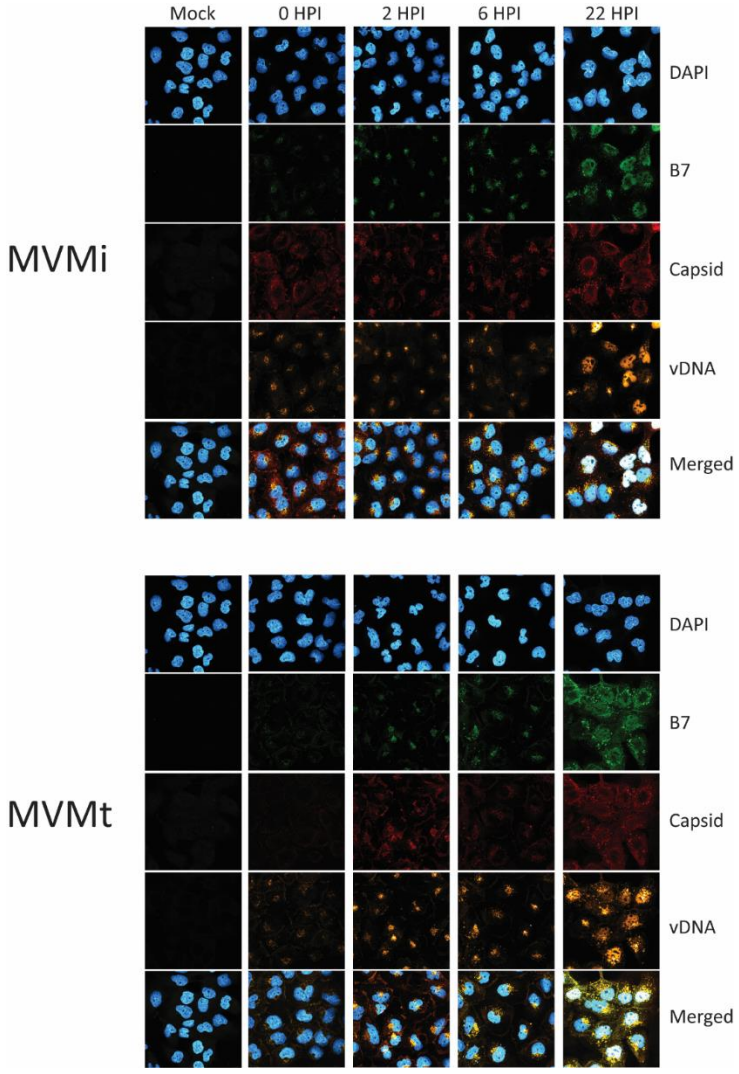


Figure 22. <u>Intracellular traffic of the MVMi and MVMt viral particles.</u>
NB324K cells were inoculated with normalized amounts (16 UHA/10e5 cells) of MVMi (equivalent to MOI 5) and T-mutant (equivalent to MOI 500) viruses. The figure shows confocal analysis of the intracellular accumulation of incoming MVMi and MVMt viruses as stained with intact-capsid (B7-Mab), capsid general epitopes (anti-MVM) and viral DNA (FISH). Representative fields of cells at the indicated hours post-infection (HPI) are shown. DAPI represents cellular nuclear DNA.

Interestingly though, neither staining of capsid proteins nor viral DNA was detected inside the nucleus at any early time points (0-6 hpi). Rather, MVMi and MVMt macromolecular signals accumulated at the nuclear periphery without significant nuclear invasion. Given the high specific infectivity of the MVMt virus (see above), these results may suggest a saturation of the cellular transport system by the amount of viral particles used in this set of experiments. Lower amounts of input virus however did not allow clear IF staining (data not shown).

7. Comparative study of PROTECTED DATA

The differential physicochemical properties and endosomal trafficking of MVMi/MVMT viruses were translated into a comparative study on the interaction of these viruses with human cancer cells. This study was aimed at assessing whether these properties increase their oncolytic capacity in culture. For this, PROTECTED DATA

Section B: Endogenization of some members of the *Parvoviridae* in primate genomes.

1. *In silico* evidences of ESVEs in primate genomes.

Endogenous parvovirus elements are widely distributed in vertebrate and invertebrate genomes (Arriagada and Gifford, 2014; Belyi *et al.*, 2010; Kapoor *et al.*, 2010; Katzourakis and Gifford, 2010; Liu *et al.*, 2011), but their presence in primate genomes has not been studied in depth. We thus screened the primate genomes available in databases for endogenous single-stranded DNA virus elements (ESVEs). The initial search was performed with eight representative members (MVM, AAV2, BPV, PBoV4-1, B19, PARV-4, AMDV-G and GmDNV), but it was subsequently extended to all members of the *Parvoviridae,* with a similar outcome. As shown in Table 12, a remarkable number of ESVEs of variable lengths belonging to the *Parvoviridae* was found at high significance in the chromosomes of different primate species from the *Cercopiticidae*. Most primate ESVEs were significantly homologous to the Rep and Capsid genes of the Adeno Associated Virus (AAV) species belonging to the *Dependoparvovirus* genus of the *Parvoviridae*. Although no infectious entities of the *Dependoparvovirus* genus have been isolated so far in monkeys (Cotmore *et al.*, 2014), we named these endogenous parvovirus sequences, based on their sequence identity, as *Simian Dependoparvovirus Elements* (SDEs). The respective primate host, nomenclature, and accession number of all the SDEs identified in this work are summarized in Table 13.

Table 12. Bioinformatic results in *Primates*

General Information				BLAST Results				
Species	Common name	Chromosome/Scaffold/Contig (genomic positions)	ID (database)	Size (bp)	Programme	Genome/Protein	E-value	Ident
Adeno-Associated Virus 2 (AAV2)								
Papio anubis	Olive Baboon	Chromosome 3 (331923 - 334302)	NW_003872382.1 (refeseq_genomics)	2375	BLASTn	Genome	0.0	84%
				1565	tBLASTn	REP68	0.0	91%
				1841		REP78	0.0	91%
				914		REP40	0.0	90%
				1191		REP52	0.0	89%
				518		VP1	4e-73	76%
				102		VP2	6e-04	71%
Cercocebus atys	Sooty Mangabey	Scaffold42 (1691316 - 1693695)	NW_012004166.1 (refeseq_genomics)	2375	BLASTn	Genome	0.0	84%
				1565	tBLASTn	REP68	0.0	89%
				1841		REP78	0.0	89%
				914		REP40	9e178	88%
				1191		REP52	0.0	88%
				518		VP1	2e-71	75%
				102		VP2	6e-04	71%
Papio anubis	Olive Baboon	Chromosome 11 (31931 -33572)	NW_003876533.1 (refeseq_genomics)	1640	BLASTn	Genome	0.0	84%
				1367	tBLASTn	REP68	0.0	92%
				1367		REP78	0.0	92%
				695		REP40	1e-139	92%
				695		REP52	1e-137	91%
Cercocebus atys	Sooty Mangabey	Scaffold60 (2893791 - 2895432)	NW_012006154.1 (refeseq_genomics)	1640	BLASTn	Genome	0.0	84%
				1367	tBLASTn	REP68	0.0	92%
				1367		REP78	0.0	92%
				695		REP40	3e-142	92%
				695		REP52	2e-140	92%
Mandrillus leucophaeus	Drill	Scaffold118 (1005600 - 1007239)	NW_012100098.1 (refeseq_genomics)	1640	BLASTn	Genome	0.0	84%
				1367	tBLASTn	REP68	0.0	90%
				1367		REP78	0.0	90%
				695		REP40	4e-141	92%
				695		REP52	5e-139	92%
Colobus angolensis	White and Black Colobus	Scaffold 190 (2545921 - 2546615)	NW_012115044.1 (refeseq_genomics)	849	BLASTn	Genome	0.0	88%
				458	tBLASTn	REP68	1e-86	86%
				458		REP78	9e-86	93%
Chlorocebus sabaeus	African Green Monkey	Chromosome 1 (18697380 - 18698248)	NW_006735234.1 (refeseq_genomics)	694	BLASTn	Genome	0.0	80%
				752	tBLASTn	REP68	1e-139	86%
				752		REP78	1e-138	86%
Torque Teno Virus 26 (TTV26)								
Rhinopithecus roxellana	Golden Snub-Nosed Monkey	Unplace genomic scaffold ENSRROG037119 (722 - 2161) (2731 - 2871) (665 - 2377)	NW_010820736.1 (refeseq_genomics)	1439	BLASTn	Genome	2e-05	66%
				140	tBLASTn	Hipothetical gp1	2e-07	55%
				1712	tBLASTn	Hypothetical gp2	5e-152	43%
		Unplace genomic scaffold ENSRROG044866 (3 - 669)	NW_010827978.1 (refeseq_genomics)	656	tBLASTn	Hypothetical gp2	7e-34	37%
		Unplace genomic scaffold ENSRROG037291 (32 - 433)	NW_010820899.1 (refeseq_genomics)	401	tBLASTn	Hypothetical gp2	2e-33	43%

Table 13. Nomenclature of endogenous parvoviridae sequences in *Cercopithecidae*

Acronym	Formal Name	Accession #
SDE_{Cat42}	Simian Dependoparvovirus Element Cercocebus atys (scaffold42)	NW_012004166.1 (1691323 - 1693695)
SDE_{Cat60}	Simian Dependoparvovirus Element Cercocebus atys (scaffold60)	NW_012006154.1 (2893791 - 2895431)
SDE_{Cdi}	Simian Dependoparvovirus Element Cercopithecus diana	Pending EMBL
SDE_{Csa}	Simian Dependoparvovirus Element Chlorocebus sabaeus	NW_006735234.1 (18697380 - 18698248)
SDE_{Can}	Simian Dependoparvovirus Element Colobus angolensis	NW_012115044.1 (2545922 - 2546615)
SDE_{Mle}	Simian Dependoparvovirus Element Mandrillus leucophaeus	NW_012100998.1 (1005600 - 1007239)
SDE_{Pan11}	Simian Dependoparvovirus Element Papio anubis (chromosome 11)	NW_003876533.1 (31932 - 33572)
SDE_{Pan3}	Simian Dependoparvovirus Element Papio anubis (chromosome 3)	NW_003872382.1 (331923 - 334302)
SDE_{Pha11}	Simian Dependoparvovirus Element Papio hamadryas (chromosome 11)	…
SDE_{Pha3}	Simian Dependoparvovirus Element Papio hamadryas (chromosome 3)	…
SDE_{Tge}	Simian Dependoparvovirus Element Theropithecus gelada	Pending EMBL

Although not identical to the known AAVs, the percentage of amino acid identity of SDEs with the Rep and Cap proteins of the currently circulating parvoviruses was in the range of 86 - 93% for Rep78 in *Colobus angolensis* and *Chlorocebus sabaeus* respectively, and 71% for VP2 and 76% for VP1 in *Papio anubis* and *Cercocebus atys* (Table 12), which is above the criteria of species assignments by the ICTV (Cotmore *et al.*, 2014). In some cases, matches of homologies were remarkably exact (E-value 0.0). Hence, in extant primates with available assembled genomic data, the ESVEs showed drastic uneven distribution, with a high presence focused in the *Cercopithecidae.* However, other studied primate families such as new world monkeys (*Cebidae: Aotus, Callithix, Saimiri, Sapajus),* asiatic tarsier (*Tarsius*) and the great apes (*Pongo, Pan, Gorilla, Homo*) of the superfamily of the *Hominicidae,* including the *Hylobatidae and Hominidae* families, gave no matches for ESVEs.

1.1. Searching ESVEs in the *Hominidae*.

The BLAST study of ESVEs in *Hominidae*, didn't show any significant similarity. This finding was difficult to reconcile with the previous suggestion of an AAV-like insertion in the *H. sapiens* genome (Liu *et al.*, 2011). We thus explored this important issue by molecular analyses, including southern-blot and PCR analysis performed in thirty genomic DNA human samples. Firstly, primers were designed from the consensus region of endogenized Rep gene in primates and the human samples were subjected to PCR amplification. Surprisingly, two out of the thirty samples amplified the expected size. These amplicons were extracted from the gels, undergone DNA sequencing and the existence of AAV like sequence was confirmed. On the other hand, Southern-blots using AAV probes failed to detect bands of related sequences (see examples in Figure 24.A). Further studies, such as deeper bioinformatics "*in silico"* search and molecular confirmation of the endogenous nature of AAVs obtained in these two samples are under investigation.

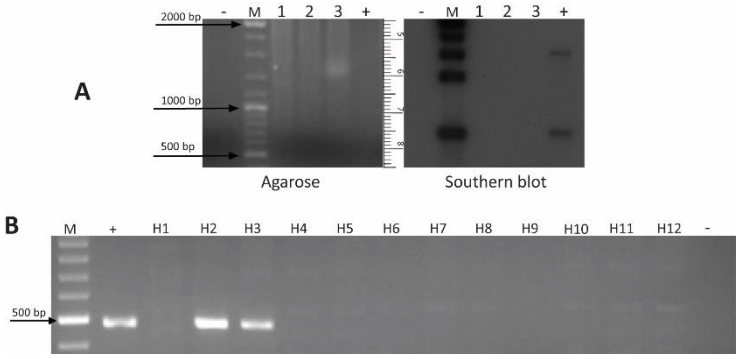


Figure 24. <u>Analysis of the presence of endogenous Parvovirus elements in human genomes.</u>
A. Southern-blot analysis of the presence of AAV sequences in humn genomes. *Left:* agarose gel electrophoresis of BamH-I restricted human DNA samples. M: molecular weight markers, 1-3: human samples, +: AAV positive control (a BamH1 digestion of the pH22 plasmid). *Right:* southern blot analysis of the same samples hybridized with a AAV ^{32}P-labeled probe. **B.** PCR analysis of the presence of AAV sequences in human samples. Primers were designed to amplify a Rep conserved region, the expected amplicon length being 434bp. Only 2 out of 30 human samples yielded the expected fragment.

1.2. Distribution and genetic structure of SDEs in the *Cercopithecidae.*

In Catarrhini monkeys, a collection of SDEs was found in different members of the *Cercopithecidae* family, which basic genetic structures are represented in Figure 25. The *Cercopithecidae* SDEs were all of highly significant nucleotide homology to the parvovirus AAV2 (Fig. 26). The coding regions within the *Cercopithecidae* SDEs consisted mainly in fragments of the Rep gene under four configurations (see Fig. 25): (a) a short Rep (1-143 amino acids), unique to *Colobus angolensis*; (b) a 1-242 aa fragment, present in *Chlorocebus sabaeus* and *Cercopithecus diana*; (c) a larger 1-445 aa fragment, shared by *Mandrillus leucophaeus, Cercocebus atys and Papio anubis* and (d) an

almost entire Rep gene (7-615 aa) plus an N-terminal fragment of about 180 aa of the capsid gene, shared by *Cercocebus atys, Papio anubis,* and *Theropithecus gelada*. Thereby the *Papio anubis* and *Cercocebus atys* primates harbored two independent insertions of SDEs in different chromosomes of their genomes.

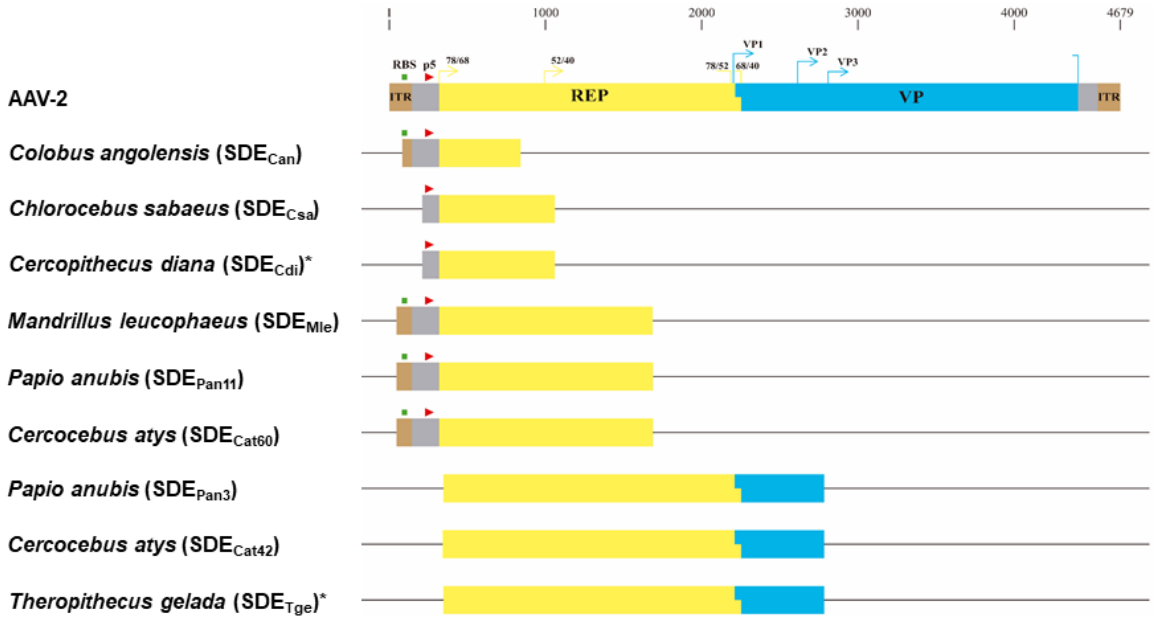


Figure 25. SDEs structures in the Cercopithecidae.
Upper: genetic organization of the AAV2 genome. The coding regions of Rep proteins and the initiation sites of the capsid proteins (VP1, VP2, VP3) are highlighted. Cis-regulatory domains, as P5 promoter and RBS (Rep binding site) within the ITR region are also illustrated. *Below:* size and genetic compositions in respect to the AAV2 genome) of the SDEs identified in the primate members of the *Cercopithecidae.*

The *Cercopithecidae* SDEs showed nucleotide homology above 80% to the currently circulating AAV2 genome (Table 12). This high conservation of nucleotide expanded uniformly across their sequences (Fig. 26) without modular heterogeneity, though the conservation was slightly higher in the REP coding regions than in the capsid gene fragment.

Of note, the SDEs showed a high degree of cis-regulatory sequence conservation. This is illustrated in Fig. 25 and 26. The three almost identical endogenous sequences (SDE$_{Cat60}$, SDE$_{Mle}$, and SDE$_{Pan11}$) remarkably maintained the B-C´-C-A-D regions, the Rep binding site (RBS), and the transcriptional regulatory sequence (TRS) of the AAV2 inverted terminal repeat (ITR) in the flip secondary structure. They do also conserve the P5 and P19 major viral promoters. A shorter A-RBS-TRS-D ITR also in the AAV2 flip configuration was found in the SDE$_{Can}$, which expanded up to include the P5 promoter.

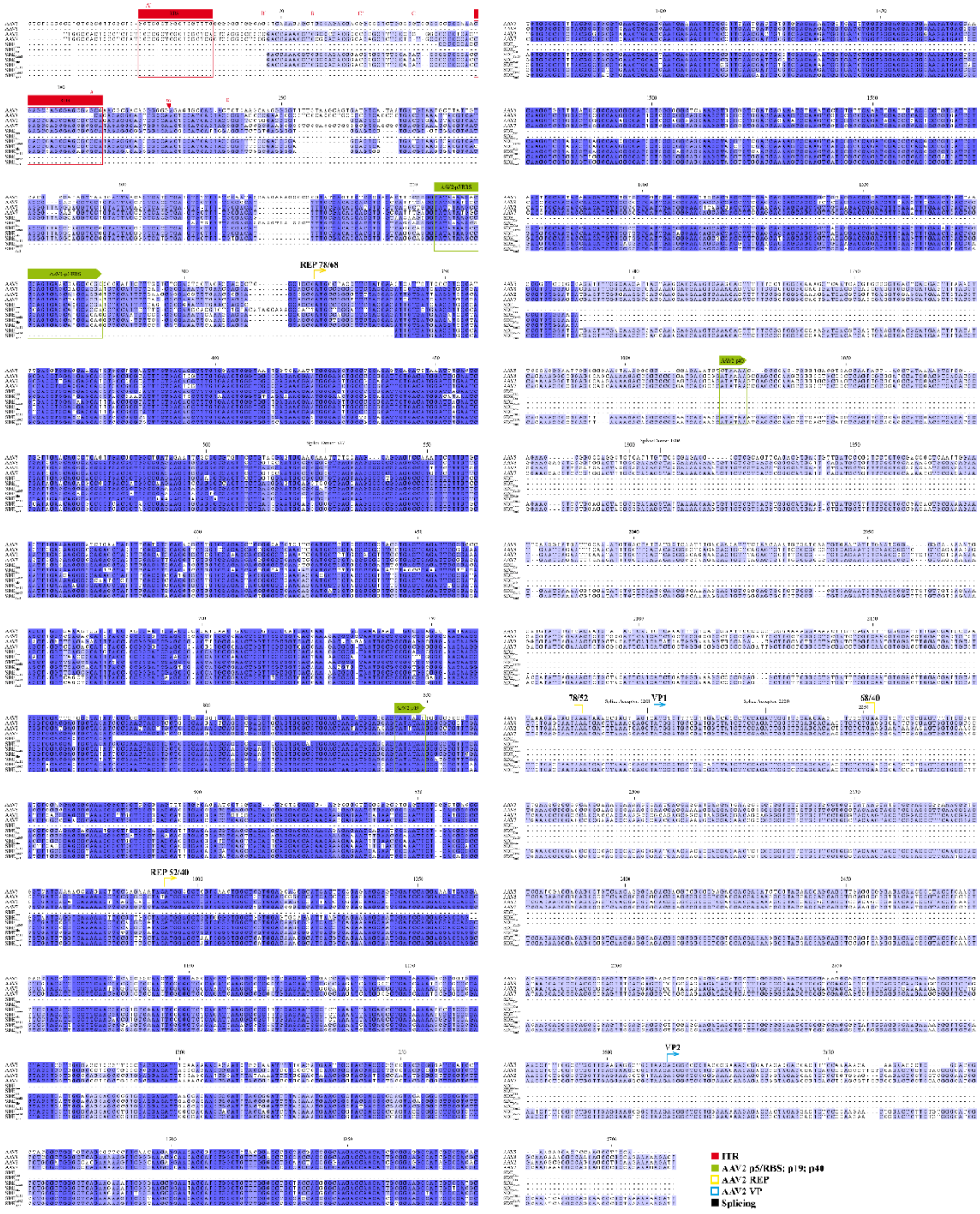


Figure 26. <u>Nucleotide alignment of some AAV genomes with the SDEs of *Cercopithecidae*.</u>
The figure shows the degree of nucleotide homology between AAV dependoparvoviruses with SDEs found in the *Cercopithecidae* primates. Some regions of the AAV2 genome, including regulatory domain, are represented. Color code is based in percentage of identity, and alignment is interrupted at the SDEs limits. Note: The general SDEs nucleotide homology are much higher compared to AAV2/7 than AAV5 sequences.

The rest of the *Cercopithecidae* SDEs lacked AAV-ITR sequences, although their composition included either the P5 (and its associated RBS) and P19 promoters ($SDE_{Csa,}$ and SDE_{Cdi}), or both P19 and P40 promoters ($SDE_{Pan3,}$ $SDE_{Cat42,}$ and SDE_{Tge}). According to figure 26, high sequence similarity was observed in most SDEs at splice donor 527 (Stutika *et al.*, 2016), and also in case of SDE_{Pan3} and SDE_{Cat42} at splice donor 1906 (Trempe and Carter, 1988). In summary, the Rep gene and the AAV2 cis-regulatory sequences showed high degree of conservation in the SDEs found in the *Cercopithecidae.*

2. Molecular confirmation of the SDEs.

To support and confirm (wherever possible) the *in silico* identified SDEs in the *Cercopithecidae* genomes and exclude possible pitfalls due to DNA contaminations (Naccache *et al.*, 2013; Smuts *et al.*, 2014), also to study some species for which no assembled sequences are available in databases yet, a large set of DNA samples obtained from primate organs were subjected to molecular analyses (Table 9). Primers were designed to amplify the different inserted SDEs, and in some cases the flanking sequences of the corresponding primate genome (Table 14). As shown in Fig. 27, the presence of SDEs corresponding to a fragment of REP sequences of the expected size was demonstrated by PCR in the genome of some *Cercopithecidae* members, as for example: *Papio hamadryas, Mandrillus sphinx, Cercocebus chrysogaster, Cercopithecus mitis* and *Theropithecus gelada*. In all cases, the amplified bands were purified from the gels and their SDE origin was confirmed by Sanger nucleotide sequencing (not shown). Some SDEs, as SDE_{Cmi}, SDE_{Tge} and SDE_{Cdi}, were found in species that had no genomic DNA available in major international nucleotide sequence databases (DNA DataBank of Japan (DDBJ), the European Nucleotide Archive (ENA), and the GenBank at NCBI). These novel SDEs were therefore submitted to ENA and Genbank databases with their unique accession numbers for further studies (Table 15). Consistently, DNA samples from primates lacking SDEs as predicted by *in silico* analysis, such as *Macaca mulatta* or *Cebus capucinus*, gave faint bands of non-corresponding sizes by PCR and upon gel-isolation their nucleotide sequences did not show homology to AAV (data not shown). Therefore, a general correspondence between the *in silico* and PCR analysis was obtained across this primate family. Interestingly, although the bioinformatic analysis predicted an SDE in *Colobus angolensis* (Fig. 25), PCR amplifications using different pairs of primers specifically designed for this SDE, as well as for flanking primate sequences, failed to detect this SDE in a sample from *Colobus polykomos,* suggesting different genome colonization by AAV between these two primate species.

Table 14. Primers used to amplifiy SDEs

Oligo Name	Seq 5'- 3'	Tm SL	bp	Expected Amp. length	Origin	Purpose
Papio-chr3-1-forward	5'-AATAGCGTAATGGGCAGGAAG-3'	54.4	21	582bp	Papio	Amplify 3' flanking region in Papio chr3, Cercocebus Scaffold42
Papio-chr3-2-reverse	5'-CAGGACGTGTAGGTGGAAATAG-3'	54.4	22			
Papio-chr3-3-forward	5'-GTTGAGGAAGTGGCTAAGACG-3'	54.7	21	640bp	Papio	Amplify 5' flanking region in Papio chr3, Cercocebus Scaffold42
Papio-chr3-4-reverse	5'-CCAGACACCGGAAAGAGGC-3'	56.5	19			
Rep-sense-AAV	5'-GTGAGTAAGGCCCCGGAGG-3'	58	19	484bp	AAV2	Amplify Rep consensus sequence in primates
Rep-antisense-AAV	5'-GCCACCCGACCAGCTCCAT-3'	61	19			
Chr11-Sense-papio	5'-TTGCGACATTTTGCGACACCA-3'	58.1	21	556bp	Papio	Amplify ITR and Rep in Papiochr11, Cercocebus scaffold 60, Mandrill
Rep-Nest-AAV2-R	5'-AGCACTCATCCACCACCTTGTT-3'	58.7	22		AAV2	
5'-Flank-Mand-F	5'-AGAAAATAGAAGCAAGACAGCAC-3'	53.6	23		Mandrill	Amplify 5' flanking region in Papio chr11, Cercocebus Scaffold60 and Mandrill
3'end Mand-F	5'-CAGCAGCCCGTGGAGGAC-3'	58.8	18	643bp	Mandrill	Detect the 3' end of Cercopithecus
3'end Mand-R	5'- GTTGTGTTTCATACCTTTTTGGCATC -3'	56.4	26			
3'end chloro-F	5'-CTACCTGCTCCCCAAAACCCA-3'	58.6	21	325bp	Chlorocebus	Detect the 3' end of Cercopithecus
3'end chloro-R	5'-TTGATGGACATTTGGGTTGGTTC -3'	56.1	23			
5-Chloro-F	5'-GTTTATTTTGTATTATGCAACCTTCC-3'	52.2	26	489bp	Chlorocebus	Amplify 5' flanking region of in Cercopithecus
5-Chloro-Cdi-R	5'-CTATCAGATTCAGGTCCATGTCAG-3'	54	24		Cercopithecus	
5'-Colobo-Flank-F	5'-TTGCTATCACCAGACCTACCTTAC-3'	56.1	24	775bp	Colobus angolensis	Amplify the whole insert in Colobus
3'-colobo-Flank-R	5'-CACTGATGAGCATTTAGGTTGATTCT-3'	56.2	26			

Table 15. List of accession numbers obtained for SDE sequences

Organism	Given Name	Accession code	Description	Provider	Release date	Database
Theropithecus gelada	SDEtge-rep	KU645401	Endogenous AAV Rep	EUPRIM-Net Project	16.11.2017	Genbank
Theropithecus gelada	SDEtge-3'	KU645402	3' insertion site of endogenous AAV	EUPRIM-Net Project	16.11.2017	Genbank
Theropithecus gelada	SDEtge-5'	ERZ481217	5' insertion site of endogenous AAV Rep	EUPRIM-Net Project	To be published	ENA
Cercopithecus aethiops (Vero Cell line)	SDEcaev-rep	ERZ496759	Endogenous AAV Rep	A. Carrascosa (ATCC)	To be published	ENA
Cercopithecus aethiops (COS1 cell line)	SDEcaec-rep	ERZ496758	Endogenous AAV Rep	A. Carrascosa (ATCC)	To be published	ENA
Cercopithecus aethiops (COS1 cell line)	SDEcaec-5'	ERZ496757	5' insertion site of endogenous AAV Rep	A. Carrascosa (ATCC)	To be published	ENA
Cercopithecus mitis monoides	SDEcmi-rep	ERZ481210	Endogenous AAV Rep	EUPRIM-Net Project	To be published	ENA
Cercopithecus mitis monoides	SDEcmi-11r	ERZ481209	5' region of AAV Rep	EUPRIM-Net Project	To be published	ENA
Cercopithecus diana	SDEcdi-11r	ERZ481208	5' region of AAV Rep	EUPRIM-Net Project	To be published	ENA
Cercopithecus diana	SDEcdi-5'	ERZ481207	5' insertion site of endogenous AAV Rep	EUPRIM-Net Project	To be published	ENA
Cercopithecus diana	SDEcdi-3'	ERZ481206	3' insertion site of endogenous AAV Rep	EUPRIM-Net Project	To be published	ENA
Papio hamadryas	SDEpha11-3'	ERZ481216	3' insertion site of endogenous AAV Rep	EUPRIM-Net Project	To be published	ENA
Papio hamadryas	SDEpha11-11r	ERZ481215	5' region of AAV Rep	EUPRIM-Net Project	To be published	ENA
Mandrillus Sphinx	SDEmsp-rep	ERZ481214	Endogenous AAV Rep	EUPRIM-Net Project	To be published	ENA
Cercocebus chrysogaster	SDEcch-rep	ERZ481205	Endogenous AAV Rep	EUPRIM-Net Project	To be published	ENA

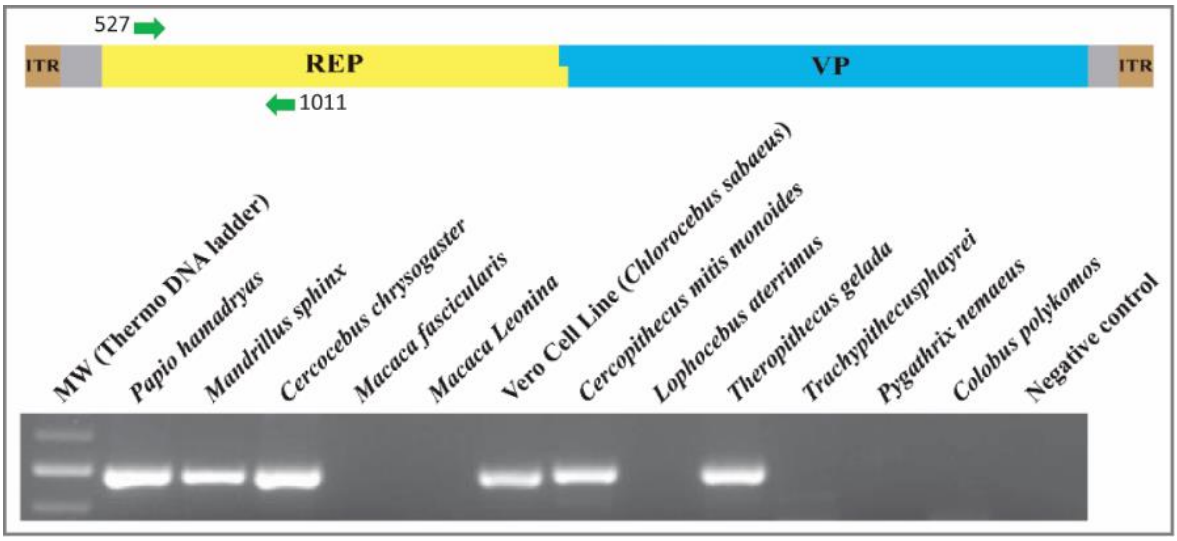


Figure 27. <u>Molecular analysis of the SDEs found in primate genomes.</u>
DNA samples obtained from the indicated primates were analyzed by PCR. Primers Rep-sense-AAV and Rep-antisense-AAV (amplifying 484bp of rep gene located between nucleotide positions 527 to 1011 of AAV genome, see Table 15) were used. The PCR products resolved by an 1.5 % agarose gel are shown. Negative control: DNA sample from human. MW, molecular weight markers (Thermofisher Scientific).

2.1. Exploration of SDEs in *Primates* with unavailable genomic data.

To extent our molecular analysis to primates of which genome sequences are not available in databases yet, we obtained DNA from the liver of *Theropoithecus gelada*, *Cercopithecus mitis*, *Lemur catta* to examine the presence of SDEs in their genomes. Southern-blotting was performed on BamHI, HindIII and EcoRI digestion products, hybridizing at high stringency with a probe containing the Rep and VP genes of AAV2 (see Materials and Methods: 4.3). The pattern and sizes of positive bands (Fig. 28, *lower*), when compared to restriction patterns expected for the phylogenetically closer reference primates genomes (Fig. 28, upper), were consistent with single insertion of a SDE of high homology to AAV2 in *Theropithecus gelada* (pattern identical to that of chromosome 3 of *P. anubis*), and other unique insertion of apparent lower homology in *Cercopithecus mitis* (pattern similar to *Chlorocebus sabaeus,* except for a HindIII fragment). These results illustrated the likeliness of findings, based on genomic sequencing, additional SDEs as it became extended to other primates.

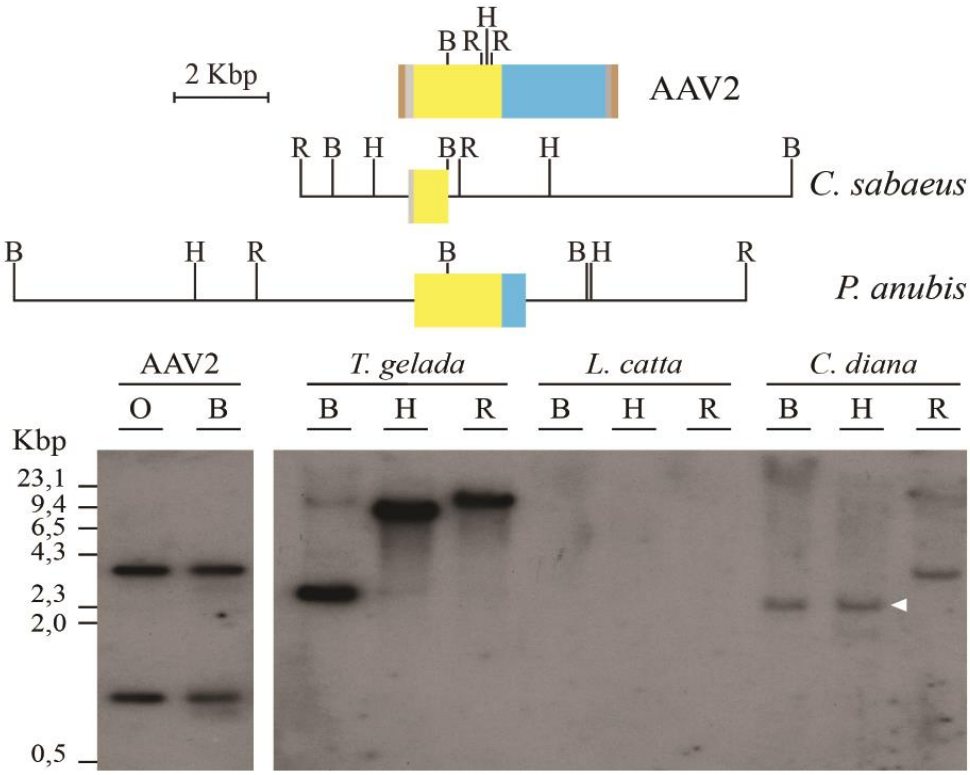


Figure 28. Southern blot analysis of SDES in primate genomes.
Upper: scheme of expected location of restriction sites in AAV2 and the reference primate genomes. *Lower:* Results obtained by the indicated primate genomes. A total of 10 µg of genomic DNA was digested with enzymes (B= BamHI; E= EcoRI and H=HindIII) and loaded per lane. Filters were hybridized with AAV probe and exposed for two weeks to intensifying screens at -70 °C. Marker track was HindIII-digested λ phage DNA that was loaded in parallel.

2.2. Study of SDEs flanking sequences in C*ercopithecidae*

Prompted by mentioned findings, we focused on molecular analysis of the SDEs insertion sites that could rule out the undesired components in the samples (like not-integrated episomal DNA, AAV viral particles of ongoing infections, or putative contaminations mentioned above) which could mislead our study on genuine endogenous parvovirus elements inserted in primate germ lines. For this, the 5´and 3´flanking regions of a large collection of the putative SDEs were amplified by PCR using specifically designed flanking primers amplifying the SDEs-primate genomes contact sequences (Table S4). As shown in Fig. 29, the resulting amplified PCR fragments fulfilled the expected sizes of the 5´and 3´flanking sequences, and furthermore were isolated from the gels and sequenced to demonstrate their endogenous nature (table new accession No). Only the SDE_{Pan11} and SDE_{Mle} 5´-flanking amplified sequences yielded an additional band of lower molecular weight of not reliable repeated sequence, in consistency with the presence of predicted

entire 5´-ITR within (Fig. 25), which commonly altered secondary structures and made the sequencing of parvovirus PCR amplification fragments difficault (unpublished observations). Moreover, this analysis allowed us to differentiate between the double SDEs insertions in the genome of Papio anubis. In summary, the consistency of *in silico* and molecular analyses of SDEs in the *Cercopithecidae* led us to conclude that their SDEs are genuine parvovirus endogenous elements that colonized the germ line of many species in this family of primates.

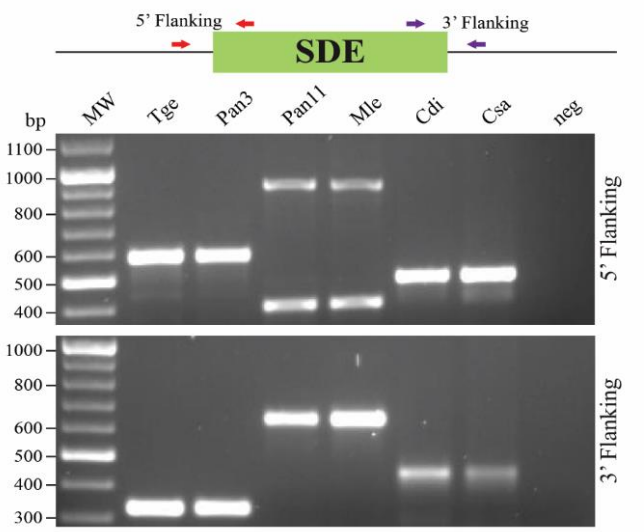


Figure 29. <u>Analysis of the integration sites of SDEs in primate genomes.</u>
PCR amplification products obtained in several members of the *Cercopithecidae* by primers designed for the 5´and 3´ SDE/primate genome junctions (see Table 15). Molecualr weight markers (Thermofisher) are shown to the left.

3. Conservation of parvovirus protein motifs in SDEs.

Parvovirus genome encodes several multifunctional non-structural proteins (mentioned as Rep or NS), and two-three structural capsid proteins (VPs). We analyzed the conservation of the best characterized parvovirus protein domains, motifs and residues across the identified SDEs sequences. Figure 30 illustrates this type of analysis focused on the rep gene, in which several SDEs of *Cercopithecidae* are compared with the circulating related AAV2, AAV7, AAV8 and the less related AAV5 parvoviruses. Two domains of Rep proteins were carefully analyzed, the Ori domain involved in parvovirus DNA binding (Ashktorab and Srivastava, 1989; Maggin *et al.*, 2012) which consists of the 100-160 amino acid residues of the protein, and the Helicase domain, also essential during virus genome replication (Walker *et al.*, 1997), encompassing the 380-450 amino acid

residues stretch. As shown in the figure, high conservation of residues between endogenous and exogenous parvovirus sequences was observed in both REP domains, remarkably in the residues essential for catalytic activities of domains, as the motifs III in the Ori domain (including two important histidine residues), and the Walker B, B´, and C of the helicase (including important residues as the two glutamic acids of walker B and the arginine in the last position of the alignment) and transactivation domain (arrows in Figure 30). This analysis demonstrated that Rep catalytic activity of viruses resulting in SDEs was proceeded through the residues analogous to the ones used by currently circulating AAVs.

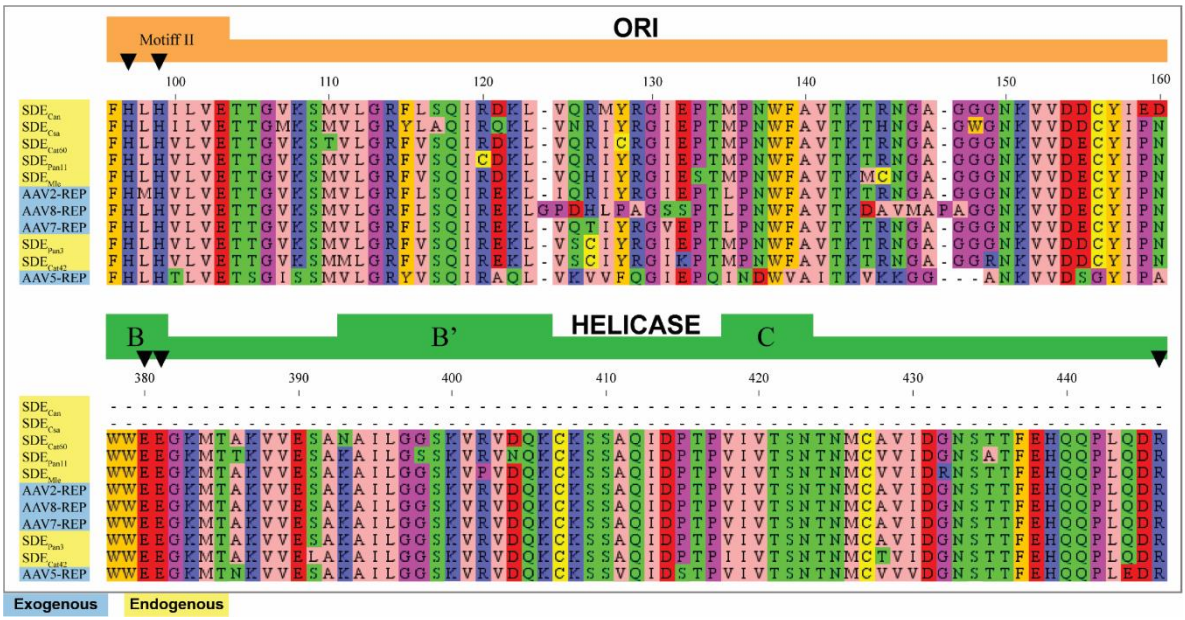


Figure 30. <u>Conservation of parvovirus Rep protein motifs in SDEs.</u>
Amino acid alignment of the REP protein of AAV parvoviruses and different SDEs identified in Primates. Blue: exogenous dependoparvoviruses, yellow: SDEs found in the *Cercopithecidae*. Conservation regions of the REP domain involved in DNA binding and helicase activity (walker A and walker B) in *Parvoviridae* are compared between AAV and the SDEs found in *Primates*. Amino acid alignment was performed by Jalview (V: 2.8.2) program and types of amino acids are colored using Clustal mode.

The alignment of SDEs with the parvovirus vp genes allowed identification of several conserved motifs characteristic to the structural proteins of *Parvoviridae.* For example, significant conservation of amino acids between SDEs and parvovirus structural proteins was found within the VP1 unique region, where cluster of basic residues are involved in virus traffic to the nucleus during cell entry (Lombardo *et al.*, 2002; Tu *et al.*, 2015; Vihinen-Ranta *et al.*, 1997), (data not shown). The conservation of residues with important functional motifs of the VP proteins is shown

in Figure 31. The N-terminal VP1 unique region contains a phospholipase domain (PLP_2) that is essential for parvovirus cell entry (Farr *et al.*, 2005; Zádori *et al.*, 2001). This PLP motif is remarkably present in the SDEs of *Cercopithecidae.* These SDEs domains show high homology with the PLP motifs of AAV, including the key residues involved in Ca^{++} binding and catalytic domains (Fig. 31, *upper*). However, the homology was much lower to best characterized phospholipases of human, primates and other origins (Fig. 31, *lower*). Therefore, this alignment also confirmed that the SDEs are of viral origin.

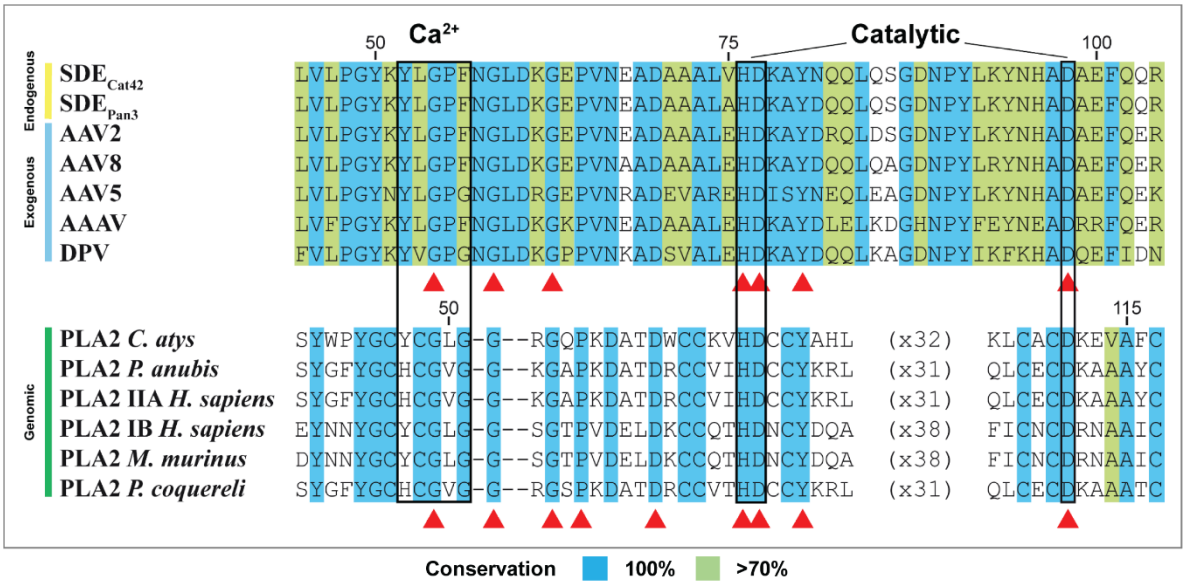


Figure 31. <u>Conservation of the PLP_2 domain in SDEs.</u>
Upper: Amino acid alignment of PLP motifs of SDEs (yellow) compared to circulating AAV (blue), as well as (*lower*) human and primate genomics (green) phospholipases. The figure illustrates the overall conservation throughout the PLP2 motif among the viral sequences, and particularly at key residues (red triangles) involved in Ca^{++} binding and catalytic functions (based on Zádori *et al.*, 2001).

4. The SDEs in the *Parvoviridae* phylogeny.

As explained above, the SDEs identified in living primates correspond to viral elements homologous to the members of *Dependoparvovirus* genus of *Parvoviridae*, which have not been isolated as infectious entities. To perform a phylogenetic analysis of the SDEs within the current taxonomy of *Parvoviridae* (Cotmore *et al.*, 2014), phylogenetic trees were constructed relating genomic fragments of the SDEs identified in primates with the homologous corresponding sequences of the rep gene of current exogenously circulating parvoviruses. As shown in Figure 32,

all Rep sequences of SDEs belonging to *Cercopithecidae*, phylogenetically clustered close to the AAV2 (member of the *A* species of AAV), the best characterized dependoparvovirus that is frequently found in human and monkey tissues. Other common dependoparvoviruses as AAV7 and AAV8 are also phylogenetically close to SDEs, whereas AAV5 and bird dependoparvoviruses are far related. Interestingly, the two SDEs found in Papio resulted in different clades: Pan3 next to Cat42, and Pan11 next to Cat60 and Mle, in consistency with their basic genetic structure (Figure 25).

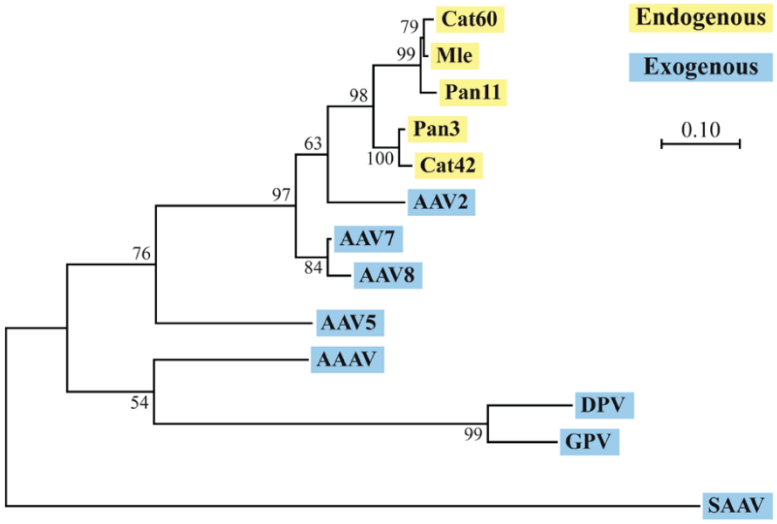


Figure 32. <u>SDEs phylogeny.</u>
Phylogenetic relationships between the SDEs (yellow) and some currently circulating Dependoparvoviruses (blue). Avian and reptile dependoviruses were used as outgroups. The program MEGA7 was used to generate the phylogenetic tree, using the maximum likelihood method with 500 bootstraps.

Discussion

Discussion

Section A:

1.Viral entry

The successful virus entry pathway leading to infection involves complex series of interactions, from the cell surface receptor to the site of virus genome uncoating in a host cell compartment. Incoming virus particles must therefore traffic through multiple membranes and vesicles acting as cellular barriers, accounting for the low specific infectivity (Si) of viral systems (Ranging from 1 to as high as 10e5 particles per infectious entity; e.g. Carpenter *et al.*, 2009), and the characteristic host cell dependence (Flint *et al.*, 2008).

In the viral system analyzed in this work, the MVMt, a Threonine mutant constructed at non-structural proteins of the immunosuppressive strain of minute virus of mice (MVMi), showed a Si more than two hundred times higher than that of the wt strain. This phenomenon was independent of the method used to measure physical virus particles (HA, TEM, or virus genomes), and the viral stocks were prepared in the same cell host along several years of research (Figure 9). This remarkably high Si, which to our knowledge is an unprecedented property for an engineered mutant virus, prompted us to address the mechanism(s) involved. Therefore, both the physical and chemical properties of viral capsids, and their early intracellular trafficking, were studied.

A major concern in our analysis was to assess whether the increased Si observed in MVMt was due to NS protein functions, or it was mediated by the viral particles. Control transfections with the MVMi and MVMt infectious genomic plasmids showed similar viral genome replication levels in NB324K cells (under elaboration), suggesting similar NS protein activities in this cell type. Moreover, the drastic physico-chemical differences of the MVMi and MVMt purified viral particles in agarose gel, under normal and stress conditions (see below), strongly suggests that their characteristic Si is mainly due to inherent viral capsid properties.

1.1. Protected Data of MVMi and MVMt capsid subunits.

A definitive evidence for different physical features between MVMi and MVMt capsids was obtained by analyzing the post-translational modifications of their protein subunits. Previous reports showed that the capsid subunits of MVMp, the prototype strain of MVM, are phosphorylated in Ser and Thr residues localized mainly at the VP2 n-terminus (Maroto *et al.*, 2000). This phosphorylation was necessary for virus egress (Maroto et al., 2004), and nuclear capsid assembly and maturation (Gil-Ranedo *et al.*, 2018; Riolobos *et al.*, 2010). On the other hand, negative charges in the viral surface were demonstrated to be important for nuclear export of MVMp, and removing of these charges by lambda phosphatase impaired the behavior of the virus in column chromatography (Wolfisberg *et al.*, 2016).

This study shows, for the first time, PROTECTED DATA

1.2. Structural and biological responses to heat of the MVMi and MVMt viral particles.

In search of properties that VP subunits phosphorylations may confer to the MVMi/t viral capsids that could explain their distinct Si, we focused on characterizing the structural and biological response of these viruses to heat *in vitro*. Highly purified empty and DNA-filled viral particles of both viruses were subjected to controlled temperature shifts (10 min) along a wide range of degrees, and then carefully inspected for integrity, configuration of capsid epitopes, behavior in native electrophoresis, hemagglutination capacity, genome uncoating and infectivity. The major changes in capsid configuration of the MVMi/t viruses in response to heat has been illustrated in

Figure 33. It should be first emphasized that, in both viruses, the infectivity strictly paralleled the degree of ssDNA genome uncoating triggered by heat, regardless to the configuration of the other studied elements of capsid surface. However, other significant differences were observed between the capsid elements of the MVMi/t viruses, as discussed below.

In TEM analysis of heat response (Figures 10 and 11), DNA-filled particles, but not empty capsids, were resistant to the penetration of uranile acetate stain in their interior. As temperature raised both viruses allowed some positive internal staining, in consistency with viral genome extrusion outside the capsid along these temperatures (Figure 15A). Thus, as reported for MVMp, the MVMi virus may tolerate some genome extrusion without capsid disassembly (Cotmore *et al.*, 2010). The major MVMi/t difference by TEM analysis was observed at high temperatures, as MVMi remained essentially intact (although mostly stained) at 80 °C, whereas the MVMt virus lost particle integrity at this temperature in accordance with genome extrusion. Therefore, the increased capsid phosphorylation of MVMt (Figure 18) may regulate a genome uncoating coupled with capsid disassembly, a property with potential significance for its high *in vivo* Si.

The analysis of virus HA capacity along temperature raise (Figure 12) also supported a higher stability of MVMi capsids than those of MVMt. Moreover, the HA study, compared to the TEM analysis, unraveled the remarkable capacity of viral capsids (mostly MVMi) to remain as intact physical particles in spite of their total loss of HA capacity. Clearly, the MVMi capsids remain assembled by subunit contacts in which HA domains are not involved. It was also interesting to find an identical full virus/empty capsid HA capacity at any temperature in both MVMi/t cases. This finding is in contrast with the stability conferred by the virus genome according to atomic force microscopy studies on MVMp (Carrasco *et al.*, 2006), suggesting that the HA domains are neither involved in MVM capsid mechanical stiffness. The HA capacity of MVMi/t viruses along temperature raise however, strictly correlated with the reactivity of electro-blotted viral capsids to anti-MVM capsid polyclonal antibody (Figure 14). These two sets of experiments highly supported the concept that HA domains and conformation epitopes on MVM capsid surface overlap and likely share amino acid residues.

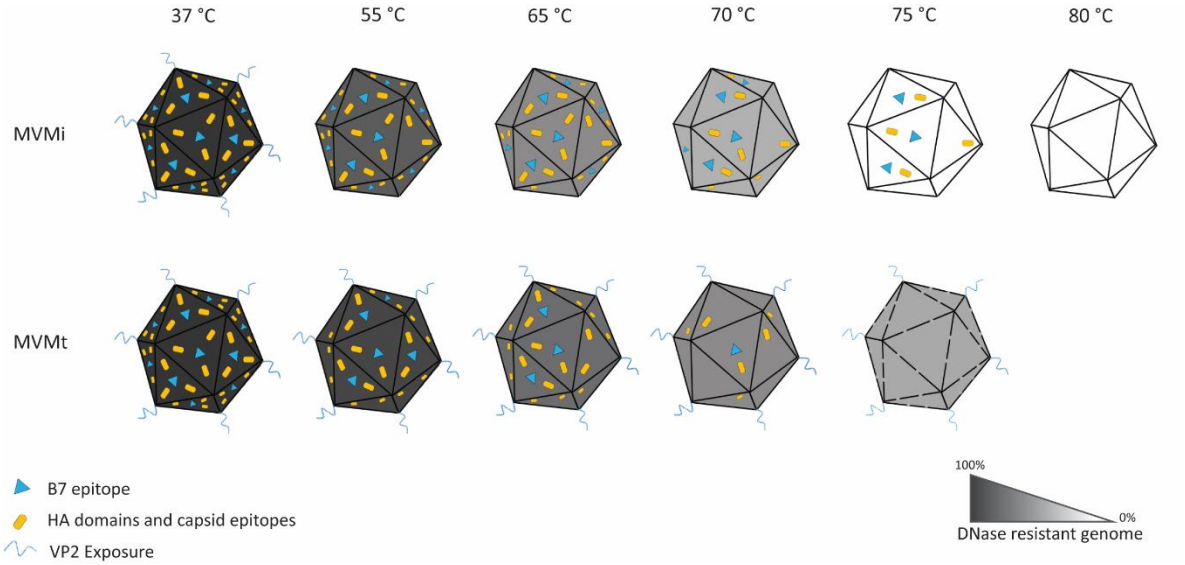


Figure 33. <u>Model of the MVMi and MVMt viral particles structural responses to heat.</u>
The figure summarizes the major differences in the heat response of the MVMi and MVMt viruses found in this study. The decrease in the number of HAU, capsid epitopes and B7 3x epitope in both MVMi/t viruses along temperature raise has been quantitatively illustrated. The exposure of the VP2 N-terminal sequence out of the capsid, and the mean extrusion of ssDNA virus genome are also illustrated. The location of the HA domains and major epitopes on the capsid surface is uncertain. Drawing of capsids reflects TEM determinations. The MVMi and MVMt capsids were not seen at 85 °C and 80 °C, respectively (Figure 11). Although the progressive minor decrease in intact virus particles as temperature raised up to 75 °C has not been quantitatively illustrated.

This study also provides novel insights into MVM virus configuration leading to capsid disassembly. MVMi/t viral particles integrity as temperature raised may be accounted by prevalence of structured capsid epitopes on the surface, including the best characterized B7-Mab epitope at the 3x axis (López-Bueno *et al.*, 2003). Importantly, the resistance of MVMi virus at high temperature was corresponded by the prevalence of the B7 reactivity, which was not observed in MVMt (Figure 14), strongly suggesting that the stability of this epitope, configured by residues of three interdigitated VP subunits (Kaufmann *et al.*, 2007), is lowered by MVMt capsid phosphorylation. Nevertheless, it is difficult to assess which VP complex catalyzes the final MVMi/t viral capsid disassembly by heat *in vitro,* because VP trimers did not conserve the B7 epitope outside of the capsid (Riolobos *et al.*, 2010). Trimers were suggested to catalyze disassembly of VLPs treated with high concentration of guanidine *in vitro* (Medrano *et al.*, 2016), although the characterization of the disassembled VP complexes was limited. Anyhow, that study is far related to our current analyses since the capsid-disassembly agent (guanidine/heat) differed, and VLPs devoid of the VP1

subunits and the phosphorylated VP2 residues that are necessary for native MVM nuclear capsid assembly in mammalian cells (Riolobos *et al.*, 2010). MVMt capsid disassembly was remarkably coupled to genome uncoating (Figure 33), therefore the putative VP complexes remain modified and should be analyzed in the context of ssDNA interaction. In MVMi the final viral stage just prior to *in vitro* disassembly is an unstable empty capsid, which is unable to withstand agarose electrophoresis. This capsid is lacking the HA domains and structured epitopes, and it is likely maintained as physical particle by the beta-barrel folding at the core of the subunits (Rossmann and Johnson, 1989). Whether this capsid entity finally disassembles as individual subunits, or some type of VP complexes, is uncertain.

The exposure out of the coat of the VP2 N-terminal domain (2Nt) was also a clearly distinct marker of the different MVMi/t heat-induced structural responses. Previous reports showed that native empty capsids, as well as VLPs of MVMp, irreversibly exposed 2Nt in respond to heat since 50-60 °C to higher temperatures (Carreira *et al.*, 2004; Hernando *et al.*, 2000; Riolobos *et al.*, 2010). However, DNA-filled MVMp virus failed to do so as the infectivity was inactivated at high temperatures (Sánchez-Martínez *et al.*, 2012). In the present report the analysis on 2Nt exposure however was complemented by additional comprehensive studies on capsid configurations, which showed that in MVMi, although 2Nt exposure is blocked at high temperatures as well (figure 16), the infectivity rather correlated with genome uncoating (Figure 15). Another important observation of the present study is the dissociation of 2Nt exposure with the overall heat-induced capsid configuration shift, as major loss of capsid epitopes and HA domains occurred at temperatures above 60 °C when the 2Nt exposure is blocked (Figure 33). These data dissect the phenomenon of MVM capsid structural transitions as a succession of events much more complex than the unique structural transition through the 50-60 °C interval previously suggested by tryptophan fluorescence (Castellanos *et al.*, 2013). Most important is the observation on VP2/3 cleavage of MVMt described in Figure 16, showing that the 2Nt exposure in this virus is highly efficient and temperature independent. These data further support that VP phosphorylation confers an increased structural plasticity to the MVMt viral capsid, which may bring advantageous properties for intracellular virus traffic in the infection.

1.3. Role of the low endosomal pH in MVMi and MVMt cell entry capacity.

Interestingly, some of our preliminary *in vitro* studies suggested that pH may severely affect MVMi/t infection properties. As observed in Figure 17A, the low pH prevented MVMi genome uncoating in response to heat, but it greatly benefited its Si at the same temperatures (Figure 17B). Remarkably, this pH beneficial effect on infection was not apparent for MVMt virions treated under the same conditions (Figure 17). The significance of this observation for the infection is difficult to assess, but it may be speculated that MVMi is over-stabilized at low pH that eventually prevents structural transitions of the incoming virions necessary to infect.

A distinct pH effect may also sustain some of our observations performed in the MVMi/t natural infection. It was reported that MVMp requires low endosomal pH for the infection (Mani *et al.*, 2006; Ros *et al.*, 2002). This requirement was also observed for MVMi here in the same cell type (NB324K cells), as drugs that raise endosomal pH (NH4Cl, or Chloroquine) did drastically inhibit MVMi infection measured by NS1 expressing cells (Figure 19). Of note, the inhibition of virus gene expression by the endosomal drugs, in infections performed at the same multiplicity (PFU/cell), was significantly higher at certain concentration for MVMt than for MVMi (Figure 19). Although still preliminary, these data may suggest that the high Si of MVMt requires low pH for efficient trafficking along the endosomal compartments. This hypothesis is currently under investigation with additional drugs and assays, particularly whether the endosomal escape may differ timely or by distinct capsid structural dynamics between both viruses.

Finally, the successful MVM infection necessarily leads to the delivery of the virus ssDNA genome into the nucleus. It would be therefore expected that the higher Si, more than two hundred folds, of MVMt in respect to MVMi results in a substantially increased nuclear delivery of the MVMt genome. We have tried to analyze the subcellular distribution of virus components during entry, but the outcome was unclear. Firstly, the genome of both viruses seemed to remain associated to the incoming capsid in a similar pattern after different hours post infection (Figure 20), suggesting that a premature MVMi uncoating may not be the key difference between these viruses. Secondly, a careful IF analysis by confocal microscopy staining capsid proteins, intact capsid, or viral DNA performed at high multiplicity of infection did not allow the detection of viral components at early time points (0-6 hpi) inside the nucleus. Rather, both MVMi and MVMt macromolecules were accumulated at the nuclear periphery without significant nuclear invasion (Figure 22). These results may suggest a saturation of the cellular transport system by the amount

of viral particles used in this set of experiments. Unfortunately, other experiments performed with lower amounts of input virus did not allow clear IF staining of viral components (data not shown). In conclusion, other more sensitive methods will be required to study the presumed differential nuclear entry of MVMi/t genomes at low multiplicity of infection.

Section B

1. Endogenous Viral Elements (EVEs)

Mostly RNA, but some DNA viruses as well, may infect germ line cells along evolution and become integrated into their chromosomes and inherited as endogenous virus elements (reviewed in Feschotte and Gilbert, 2012). Most identified endogenous ssDNA virus elements (ESVEs) belong to the *Parvoviridae* (Arriagada and Gifford, 2014; Belyi *et al.*, 2010; Kapoor *et al.*, 2010; Katzourakis and Gifford, 2010; Liu *et al.*, 2011), a family of viruses that, as explained in introduction, successfully infects hosts belonging to the three domains of life. The increasing availability of genomic sequences and annotations in data bases prompted us to search for ESVEs in primates.

2. Presence of ESVEs in non-human primate genomes.

This work describes the first in depth *in silico* analysis, supported by molecular characterizations, of genetic elements of the ssDNA viruses inserted in the germ line (ESVEs) of some currently living primates. The elements were designated as Simian Dependoparvovirus Elements (SDEs), as they reflect the insertion of uncharacterized parvoviruses of the *Dependoparvovirus* genus of the *Parvoviridae,* which must have circulated (or are still circulating) as exogenous infectious agents in primates prior endogenization. It should be emphasized that these SDEs correspond only to those genetic elements found with E-values equal or less than 1e-10, the threshold that we set as significant, but it may be possible that other SDE(s) with less significance, or genomically spread sequences with reminiscence of parvoviruses genomes be found in future searches.

Our wide search of ESVEs in primates resulted in multiple SDEs of the *Parvoviridae*. This finding supports other studies that highlighted the dominant relevance of a parvoviral origin for the

collection of ssDNA virus elements found in many prokaryotic and eukaryotic organisms (Katzourakis and Gifford, 2010). This predominance of SDEs elements is remarkable and suggests biological inherent properties of parvoviruses promoting or favouring their germinal insertion. It may be suggested that the efficient endogenization of these ssDNA viruses is related to the rolling-circle mechanism of replication (Tattersall and Ward, 1976) and their insertional mechanisms in non-productive infection (Samulski *et al.*, 1991). Moreover, other possible factors could favor the successful endogenization of parvoviruses, such as their nuclear life cycle (Cotmore and Tattersall, 2007), the capacity to persist as non-toxic infectious entities in growth-arrested cells (Gil-Ranedo *et al.*, 2015), the systemic infection and blood transmision (Qiu *et al.*, 2017), and their capacity to infect stem cells in tissues (Segovia *et al.*, 2003).

3. ESVEs insertion events throughout primate evolutionary history.

We have found ESVE of the *Parvoviridae* that has been endogenized in certain families of Primates. A summary of our collection data is illustrated in Fig. 34. No ssDNA virus endogenization were identified in the new world monkeys (Platyrrhini), neither in the *Lorisiodea (Galagidae* and *Lorisidae* families*)* or *Hominicidae* (*Hylobatidae and Hominidae* families*)* superfamilies of primates. Three families of primates, the *Atelidae*, *Pithecidae* (belonging to the new world monkeys Platyrrhini) and the Strepsirrhini were not analyzed due to unavailability of DNA samples and the lack of DNA sequence information in databases.

A total number of four independent endogenization events of ESVEs were localized in genomes of currently living primates. The precise number of SDE insertion sites could be deduced from their genetic structure and primate phylogeny. Usually a single SDE or in some cases two SDEs were found in several primate genera of the *Cercopithecidae* family. The genetic structure and phylogeny data indicated that the endogenization events leading to SDEs occurred after the split of the *Haplorrhini* and *Strepsirrhini* primates.

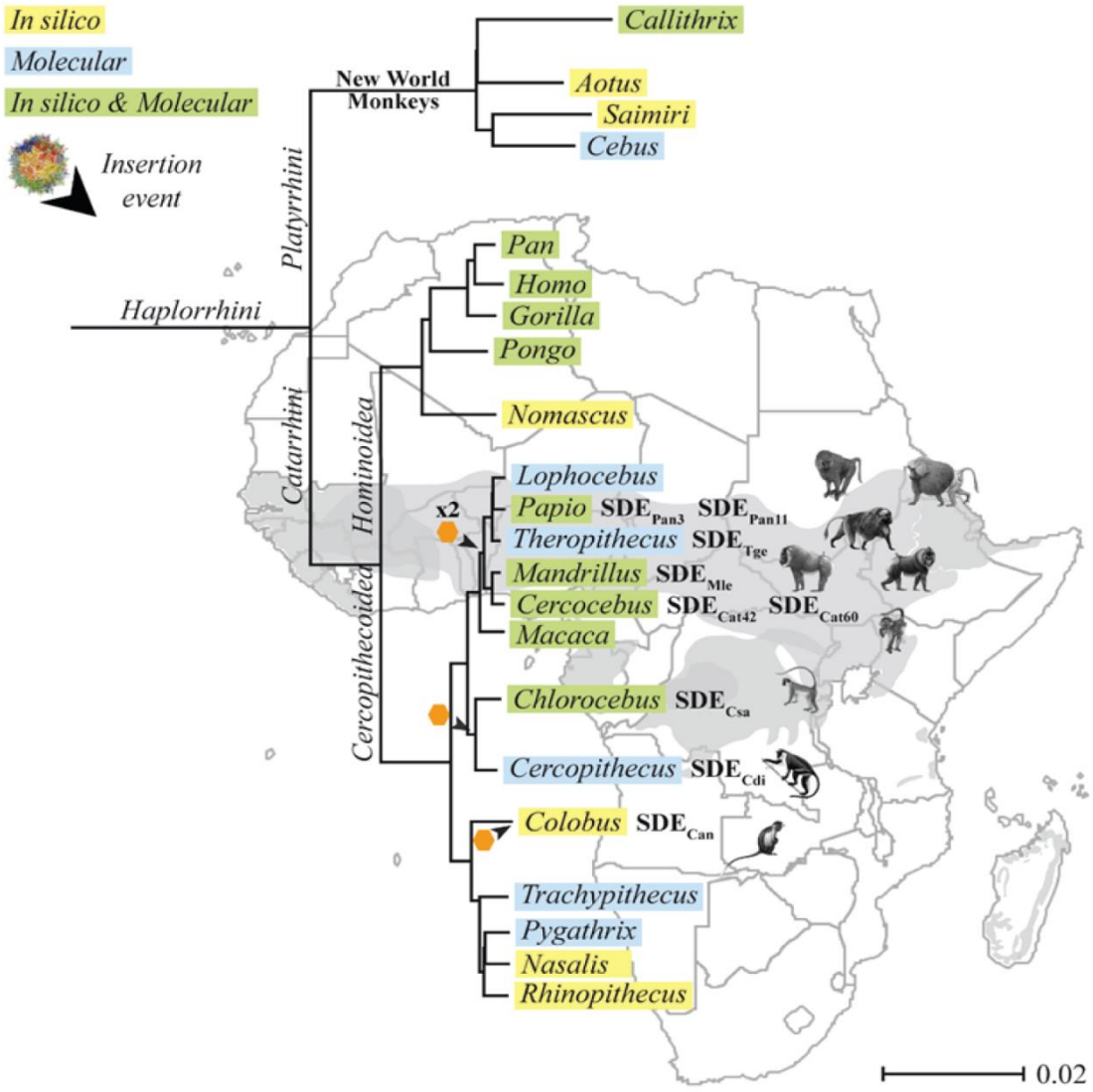


Figure 34. **Insertion events of AAV elements in the *Haplorrhini* primate genomes along evolution**
The figure illustrates the genera of the Primates phylogeny (adapted from Perelman *et al.*, 2011) tested in this report for the presence of endogeneous parvovirus elements (SDE) by *in silico* and/or molecular analyses. Arrows mark the identified endogenization events illustrated by coloured capsids at branches of primate phylogeny. The SDEs are shown in one primate species of each genus. In those primate genera carrying two SDE a single species representing each type of insertion is shown. Illustrations and approximate geographical distribution of African primates are from Petter and Desbordes, 2010.

4. The origin of SDEs

The origin of the endogenized AAV cannot be traced easily. Lower evolutionary rates of endogenous viruses in respect to circulating viruses have been decribed (Feschotte and Gilbert, 2012) and must be taken into account. Interestingly, the SDEs found in the primate species of the *Cercopithecidae* are much closer to the AAV2 parvovirus (Figure 32). Although these homologies do not imply how old the SDEs are, distinct origin(s) among them seems likely; However, they do not indicate the host range of the ancient circulating endogenized AAV since the tropism determinants of parvoviruses could rely on a few distinct amino acids (Cotmore and Tattersall, 2007; López-Bueno *et al.*, 2006; Parrish, 2010). Different *Dependoparvoviruses* may have been endogenized as the primate speciations occurred, about 40-60 million years ago, paralleling the different patterns of primate species in the main land of the African continent (*Cercopithecidae*).

5. Evolutionary implications of ESVE insertions across primate phylogeny.

Although our study included most primate genomes available in 2016/17 databases, as well as some DNA samples from primates not sequenced yet, it is possible that when the sequence of other primate genomes became available, particularly Asian *Cercopithecidae* and the new world monkeys (*Platyrrhini*), the overall presence of SDEs in the biosphere became notoriously enlarged.

Some evolutionary implications of SDEs for the dynamics of primate genomes can be hypothesized. The presence of two distinct SDEs in *Papio* and *Cercocebus,* with identical structures was confirmed by genome analysis and PCR amplifications (Figure 25-29) but the existance of only the shorter SDE in *Mandrillus leucophaeus,* cannot be accommodated with the best stablished primates phylogenetic tree, as *Mandrillus* and *Cercocebus* are closer than *Papio* (Perelman *et al.*, 2011). These observations may suggest the loss of the larger SDE locus in *Mandrillus* which would support a high dynamic and plasticity of the primate genomes.

6. Parvovirus endogenization mechanism(s)

In eukaryots, most viral endogenization mechanisms operate from RNA messengers or input RNA virus genomes, which are converted to dsDNA at certain stages of virus life cycle. It has been known for decades that the *Retroviridae* ssRNA+ genome is reverse-transcribed and integrated to initiate infection. Endogenous retrovirus elements (ERVs) may be generated through a diversity of molecular rearrangements (Feschotte and Gilbert, 2012) and can be subsequently mobilized, by

reverse-transcription activities, in thousands of copies across the invaded genome. Other RNA viruses, lacking reverse transcriptase, may be endogenized via cellular retrotransposons acting on viral RNA transcripts generated in the infection. For example, the endogenous elements of Bornaviruses were corresponded to pseudogenes flanked by repeat sequences and polyA stretches indicative of LINE-1 retrotranscription activity (Horie *et al.*, 2010). However, endogenized virus genes lacking traces of LINE-1 activity could also be identified. It was suggested that other RNA virus families might follow different endogenization pathways. The mechanism(s) of DNA virus endogenizations are largerly unknown, although it is belived that generally a RNA molecule is served as an intermediate (Katzourakis and Gifford, 2010).

This study shows that parvovirus colonization of primate germ lines likely involved a different pattern of DNA-mediated insertional mechanism(s). All the identified SDEs conserve cis-acting regulatory sequences, which are not present in the *Dependoparvovirus* transcripts. For example the 5´-ITR, P5, and P19 promoters are conserved in many SDEs of the *Cercopithecidae* (Figure 25). Also most SDEs do conserve introns. Collectively, these genetic features delineate an endogenization mechanism that is driven by the incoming ssDNA parvovirus genome, or a dsDNA replicative intermediate.

AAV replicative intermediates and other DNA forms are naturally found accumulated in mammalian tissues, notably in human and non-human primates (Gao *et al.*, 2004; Schnepp *et al.*, 2005). Also other parvovirus DNA forms may persist in experimental conditions (Segovia *et al.*, 2003), which could eventually favour a recombination process with host genome. Interestingly, an insertional mutagenesis caused by AAV elements in mice and humans lead to hepatocarcinoma (Donsante *et al.*, 2007; Nault *et al.*, 2015). The machinery responsible of SDE insertions is largerly unknown. Uncovering this machinery may help us understand why some species are better substrate for genome colonization, and also it could contribute to the development of safer *Dependoparvovirus* based vectors for human and primate gene therapy.

7. Perspectives

Clearly, further and extensive research is required prior to a comprehensive understanding of the functional role(s) that the SDEs may, or may not have, in the evolution of primates. The successful colonization of the genome of several primates families by AAV opens new perspectives for the future evolution of human and non-human primates. The prevalence of AAV circulating as

infectious entities in human (Gao *et al.*, 2004; Schnepp *et al.*, 2005) and Primates (Gao *et al.*, 2003, 2004) is very high. Although we have been unable to detect ESVEs in the available genomes of *Hominidae* (neither in a collection of human samples), the high rate of circulating AAV as well as its widely demonstrated capacity to become inserted in human cells and tissues (Hüser *et al.*, 2010; Janovitz *et al.*, 2013; Kotin *et al.*, 1992; Samulski *et al.*, 1991), and to colonize the germ line of many animals (Katzourakis and Gifford, 2010) and primates (this study), allow us to anticipate that endogenization of AAV will continue to occur, resulting in ESVE in future or currently living individual and certain human populations. A likely ongoing phenomenon which may shape the genome and perhaps contribute to future evolution of human and non-human primates.

Conclusions

Conclusions

1. An electrophoretic-blot method for viral particle analysis, set-up in this study, allows the resolution of MVM empty and DNA-filled viral particles, as well as tracing some shifts in capsid stability and configuration.

2. The VP2-Nt sequence exposure out of the capsid coat contributes to the heterogeneous electrophoretic mobility of the DNA-filled MVM virions in agarose gels.

3. The DNA-filled MVMt mutant virus exhibits more pronounced exposure dynamics of the 2Nt and 1Nt sequences than those of the MVMi wild type virus, accounting for its increased heterogeneous mobility in gels.

4. The biological and physicochemical properties of MVMi and MVMt full viruses and empty capsids respond differently to heat under neutral and acidic pH conditions.

5. PROTECTED DATA

6. The higher killing capacity of human cancer cells by MVMt in respect to that of MVMi is related, at least in part, to a differential trafficking through the endosomal network.

7. Nine genetic endogenous elements of the *Parvoviridae,* here named Simian Dependoparvovirus Elements (SDEs), were characterized in the genome of primates belonging to the *Cercopithecidae.*

8. The SDEs show a high level of nucleotide and amino acid identity with currently circulating virus members of the *Dependoparvovirus*.

9. A phylogenetic relationship has been established between the SDEs and the currently circulating *Dependoparvovirus* genomes.

10. We propose four independent invasions of primate genomes by parvoviruses along evolution.

Conclusiones

1. Un método de electroforesis en agarosa y transferencia a filtros, elaborado en este estudio, permite la resolución de partículas víricas de MVM vacías y llenas de ADN, así como detectar cambios en la estabilidad y configuración de las cápsidas.

2. La exposición de la secuencia VP2-Nt fuera de la cápsida contribuye a la movilidad electroforética heterogénea en geles de agarosa de los viriones MVM llenos de ADN.

3. El virus mutante MVMt lleno de ADN exhibe una dinámica de exposición más pronunciada de las secuencias 2Nt y 1Nt que las del virus silvestre MVMi, lo que explica la mayor heterogeneidad en su migración en geles de agarosa.

4. Las propiedades biológicas y fisicoquímicas de los virus completos MVMi y MVMt, así como de sus cápsidas vacías, responden de manera diferente al calor en condiciones de pH neutro y ácido.

5. DATOS PROTEGIDOS

6. El incremento en la capacidad de matar células cancerosas humanas por el virus MVMt, con respecto a la del MVMi, está relacionada, al menos en parte, con un tráfico diferente a través de la red endosomal.

7. Se han caracterizado nueve elementos genéticos endógenos de *Parvoviridae*, aquí llamados Elementos de Dependoparvovirus de simio (EDS, o SDE), en el genoma de primates pertenecientes a la familia *Cercopithecidae*.

8. Los SDE muestran un alto nivel de identidad de nucleótidos y aminoácidos con algunos virus circulantes en la actualidad del género *Dependoparvovirus.*

9. Se ha establecido una relación filogenética entre los SDE y los genomas de los Dependoparvoviruses actualmente circulantes.

10. Proponemos cuatro invasiones independientes de genomas de Primates por parvovirus a lo largo de la evolución.

گفتند یافت می‌نشود گشته‌ایم ما

گفت آنک یافت می‌نشود آنم آرزوست

مولانا، (۶۰۴ الی ۶۷۲ قمری)

Dedicated to Pepe:

They said, “It is not to be found; we have searched too.”
He answered, “the thing who is not to be found is my desire.”

Dijeron, "No se encuentra, nosotros también lo hemos buscado".
Respondió, "lo que no se encuentra es mi deseo".

Rumi (Molana), Persian Poet (1207- 1273)

References

References

1. Agbandje-McKenna, M., Llamas-Saiz, a L., Wang, F., Tattersall, P. and Rossmann, M. G. (1998) Functional implications of the structure of the murine parvovirus, minute virus of mice., *Structure*, 6 (11), pp. 1369–81. DOI:10.1016/S0969-2126(98)00137-3.

2. Agola, J. O., Jim, P. A., Ward, H. H., Basuray, S. and Wandinger-Ness, A. (2011) Rab GTPases as regulators of endocytosis, targets of disease and therapeutic opportunities., *Clinical Genetics*, 80 (4), pp. 305–18. DOI:10.1111/j.1399-0004.2011.01724.x.

3. Allaume, X., El-Andaloussi, N., Leuchs, B., Bonifati, S., Kulkarni, A., Marttila, T., *et al.* (2012) Retargeting of rat parvovirus H-1PV to cancer cells through genetic engineering of the viral capsid., *Journal of Virology*, 86 (7), pp. 3452–65. DOI:10.1128/JVI.06208-11.

4. Allison, A. B., Kohler, D. J., Ortega, A., Hoover, E. A., Grove, D. M., Holmes, E. C. and Parrish, C. R. (2014) Host-specific parvovirus evolution in nature is recapitulated by in vitro adaptation to different carnivore species., *PLoS Pathogens*, 10 (11), pp. e1004475. DOI:10.1371/journal.ppat.1004475.

5. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) Basic local alignment search tool, *Journal of Molecular Biology*, 215 (3), pp. 403–410. DOI:10.1016/S0022-2836(05)80360-2.

6. Angelova, A. L., Geletneky, K., Nüesch, J. P. F. and Rommelaere, J. (2015) Tumor Selectivity of Oncolytic Parvoviruses: From in vitro and Animal Models to Cancer Patients., 3, pp. 55. DOI:10.3389/fbioe.2015.00055.

7. Antonietti, J. P., Sahli, R., Beard, P. and Hirt, B. (1988) Characterization of the cell type-specific determinant in the genome of minute virus of mice., *Journal of Virology*, 62 (2), pp. 552–7.

References

8. Arbuckle, J. H., Medveczky, M. M., Luka, J., Hadley, S. H., Luegmayr, A., Ablashi, D., *et al.* (2010) The latent human herpesvirus-6A genome specifically integrates in telomeres of human chromosomes in vivo and in vitro., *Proceedings of the National Academy of Sciences of the United States of America*, 107 (12), pp. 5563–8. DOI:10.1073/pnas.0913586107.

9. Arriagada, G. and Gifford, R. J. (2014) Parvovirus-Derived Endogenous Viral Elements in Two South American Rodent Genomes, *Journal of Virology*, 88 (20), pp. 12158–12162. DOI:10.1128/JVI.01173-14.

10. Ashktorab, H. and Srivastava, A. (1989) Identification of nuclear proteins that specifically interact with adeno-associated virus type 2 inverted terminal repeat hairpin DNA., *Journal of Virology*, 63 (7), pp. 3034–9.

11. Balazs, A. B., Chen, J., Hong, C. M., Rao, D. S., Yang, L. and Baltimore, D. (2012) Antibody-based protection against HIV infection by vectored immunoprophylaxis, *Nature*, 481 (7379), pp. 81–84. DOI:10.1038/nature10660.

12. Ball-Goodrich, L. J. and Tattersall, P. (1992) Two amino acid substitutions within the capsid are coordinately required for acquisition of fibrotropism by the lymphotropic strain of minute virus of mice., *Journal of Virology*, 66 (6), pp. 3415–23.

13. Basak, S. and Turner, H. (1992) Infectious entry pathway for canine parvovirus., *Virology*, 186 (2), pp. 368–76. DOI:10.1016/0042-6822(92)90002-7.

14. Bashir, T., Horlein, R., Rommelaere, J. and Willwand, K. (2000) Cyclin A activates the DNA polymerase delta -dependent elongation machinery in vitro: A parvovirus DNA replication model., *Proceedings of the National Academy of Sciences of the United States of America*, 97 (10), pp. 5522–7. DOI:10.1073/pnas.090485297.

References

15. Bejarano, E. R., Khashoggi, A., Witty, M. and Lichtenstein, C. (1996) Integration of multiple repeats of geminiviral DNA into the nuclear genome of tobacco during evolution., *Proceedings of the National Academy of Sciences*, 93 (2), pp. 759–764. DOI:10.1073/pnas.93.2.759.

16. Belshaw, R. and Katzourakis, A. (2005) BlastAlign: a program that uses blast to align problematic nucleotide sequences, *Bioinformatics*, 21 (1), pp. 122–123. DOI:10.1093/bioinformatics/bth459.

17. Belyi, V. A., Levine, A. J. and Skalka, A. M. (2010) Sequences from Ancestral Single-Stranded DNA Viruses in Vertebrate Genomes: the Parvoviridae and Circoviridae Are More than 40 to 50 Million Years Old, *Journal of Virology*, 84 (23), pp. 12458–12462. DOI:10.1128/JVI.01789-10.

18. Berns, K. I. and Linden, R. M. (1995, March) The cryptic life style of adenoassociated virus, *BioEssays*. DOI:10.1002/bies.950170310.

19. Boisvert, M., Bouchard-Lévesque, V., Fernandes, S. and Tijssen, P. (2014) Classic nuclear localization signals and a novel nuclear localization motif are required for nuclear transport of porcine parvovirus capsid proteins., *Journal of Virology*, 88 (20), pp. 11748–59. DOI:10.1128/JVI.01717-14.

20. Boisvert, M., Fernandes, S. and Tijssen, P. (2010) Multiple pathways involved in porcine parvovirus cellular entry and trafficking toward the nucleus., *Journal of Virology*, 84 (15), pp. 7782–92. DOI:10.1128/JVI.00479-10.

21. Bonnard, G. D., Manders, E. K., Campbell, D. A., Herberman, R. B. and Collins, M. J. (1976) Immunosuppressive activity of a subline of the mouse EL-4 Lymphoma, *J Exp Med*, 143 (1), pp. 187–205. DOI:10.1084/jem.143.1.187.

References

22. Bowman, E. J., Siebers, A. and Altendorf, K. (1988) Bafilomycins: a class of inhibitors of membrane ATPases from microorganisms, animal cells, and plant cells., *Proceedings of the National Academy of Sciences of the United States of America*, 85 (21), pp. 7972–7976. DOI:10.1073/pnas.85.21.7972.

23. Brownstein, D. G., Smith, A. L., Jacoby, R. O., Johnson, E. A., Hansen, G. and Tattersall, P. (1991) Pathogenesis of infection with a virulent allotropic variant of minute virus of mice and regulation by host genotype., *Laboratory Investigation; a Journal of Technical Methods and Pathology*, 65 (3), pp. 357–64.

24. Buchholz, C. J., Friedel, T. and Büning, H. (2015) Surface-Engineered Viral Vectors for Selective and Cell Type-Specific Gene Delivery., *Trends in Biotechnology*, 33 (12), pp. 777–790. DOI:10.1016/j.tibtech.2015.09.008.

25. Canaan, S., Zádori, Z., Ghomashchi, F., Bollinger, J., Sadilek, M., Moreau, M. E., Tijssen, P. and Gelb, M. H. (2004) Interfacial enzymology of parvovirus phospholipases A2., *The Journal of Biological Chemistry*, 279 (15), pp. 14502–8. DOI:10.1074/jbc.M312630200.

26. Carpenter, J. E., Henderson, E. P. and Grose, C. (2009) Enumeration of an extremely high particle-to-PFU ratio for Varicella-zoster virus., *Journal of Virology*, 83 (13), pp. 6917–21. DOI:10.1128/JVI.00081-09.

27. Carrasco, C., Carreira, A., Schaap, I. A. T., Serena, P. A., Gomez-Herrero, J., Mateu, M. G. and de Pablo, P. J. (2006) DNA-mediated anisotropic mechanical reinforcement of a virus, *Proceedings of the National Academy of Sciences*, 103 (37), pp. 13706–13711. DOI:10.1073/pnas.0601881103.

28. Carreira, A., Menéndez, M., Reguera, J., Almendral, J. M. and Mateu, M. G. (2004) In Vitro Disassembly of a Parvovirus Capsid and Effect on Capsid Stability of Heterologous Peptide Insertions in Surface Loops, *Journal of Biological Chemistry*, 279 (8), pp. 6517–6525. DOI:10.1074/jbc.M307662200.

References

29. Castellanos, M., Pérez, R., Rodríguez-Huete, A., Grueso, E., Almendral, J. M. M. and Mateu, M. G. G. (2013) A slender tract of glycine residues is required for translocation of the VP2 protein N-terminal domain through the parvovirus MVM capsid channel to initiate infection., *Biochemical Journal*, 455 (1), pp. 87–94. DOI:10.1042/BJ20130503.

30. Chang, S. F., Sgro, J. Y. and Parrish, C. R. (1992) Multiple amino acids in the capsid structure of canine parvovirus coordinately determine the canine host range and specific antigenic and hemagglutination properties., *Journal of Virology*, 66 (12), pp. 6858–67.

31. Chuong, E. B., Elde, N. C. and Feschotte, C. (2016) Regulatory evolution of innate immunity through co-option of endogenous retroviruses., *Science (New York, N.Y.)*, 351 (6277), pp. 1083–7. DOI:10.1126/science.aad5497.

32. Clamp, M., Cuff, J., Searle, S. M. and Barton, G. J. (2004) The Jalview Java alignment editor, *Bioinformatics*, 20 (3), pp. 426–427. DOI:10.1093/bioinformatics/btg430.

33. Cohen, S., Behzad, A. R., Carroll, J. B., Panté, N. and Pante, N. (2006) Parvoviral nuclear import: Bypassing the host nuclear-transport machinery, *Journal of General Virology*, 87 (11), pp. 3209–3213. DOI:10.1099/vir.0.82232-0.

34. Cotmore, S. F., Agbandje-McKenna, M., Chiorini, J. A., Mukha, D. V., Pintel, D. J., Qiu, J., *et al.* (2014) The family Parvoviridae, *Archives of Virology*, 159 (5), pp. 1239–1247. DOI:10.1007/s00705-013-1914-1.

35. Cotmore, S. F., D'Abramo, A. M., Carbonell, L. F., Bratton, J. and Tattersall, P. (1997) The NS2 polypeptide of parvovirus MVM is required for capsid assembly in murine cells, *Virology*, 231 (2), pp. 267–280. DOI:10.1006/viro.1997.8545.

36. Cotmore, S. F., D'Abramo, A. M., Ticknor, C. M. and Tattersall, P. (1999) Controlled Conformational Transitions in the MVM Virion Expose the VP1 N-Terminus and Viral Genome without Particle Disassembly, *Virology*, 254 (1), pp. 169–181. DOI:10.1006/viro.1998.9520.

References

37. Cotmore, S. F., Hafenstein, S. and Tattersall, P. (2010) Depletion of virion-associated divalent cations induces parvovirus minute virus of mice to eject its genome in a 3'-to-5' direction from an otherwise intact viral particle., *Journal of Virology*, 84 (4), pp. 1945–56. DOI:10.1128/JVI.01563-09.

38. Cotmore, S. F. and Tattersall*, P. (1995) DNA replication in the autonomous parvoviruses, *Seminars in Virology*, 6 (5), pp. 271–281. DOI:10.1006/SMVY.1995.0033.

39. Cotmore, S. F. and Tattersall, P. (1987) The autonomously replicating parvoviruses of vertebrates, *Advances in Virus Research*, 33 (C), pp. 91–174. DOI:10.1016/S0065-3527(08)60317-6.

40. Cotmore, S. F. and Tattersall, P. (1989) A genome-linked copy of the NS-1 polypeptide is located on the outside of infectious parvovirus particles., *Journal of Virology*, 63 (9), pp. 3902–11.

41. Cotmore, S. F. and Tattersall, P. (2007) Parvoviral host range and cell entry mechanisms., *Advances in Virus Research*, 70, pp. 183–232. DOI:10.1016/S0065-3527(07)70005-2.

42. Crawford, L. (1966) A minute virus of mice, *Virology*, 29, pp. 605–612. DOI:10.1016/0042-6822(66)90284-4.

43. Crochu, S., Cook, S., Attoui, H., Charrel, R. N., De Chesse, R., Belhouchet, M., Lemasson, J.-J., de Micco, P. and de Lamballerie, X. (2004) Sequences of flavivirus-related RNA viruses persist in DNA form integrated in the genome of Aedes spp. mosquitoes, *Journal of General Virology*, 85 (7), pp. 1971–1980. DOI:10.1099/vir.0.79850-0.

44. Deleu, L., Baldauf, A. Q., Rommelaere, J., Beard, P., Willwand, K., Costello, E. and Mumtsidu, E. (1997) The minute virus of mice (MVM) nonstructural protein NS1 induces nicking of MVM DNA at a unique site of the right-end telomere in both hairpin and duplex conformations in vitro., *Journal of General Virology*, 78 (10), pp. 2647–2655. DOI:10.1099/0022-1317-78-10-2647.

References

45. Diering, G. H. and Numata, M. (2014) Endosomal pH in neuronal signaling and synaptic transmission: role of Na(+)/H(+) exchanger NHE5., *Frontiers in Physiology*, 4, pp. 412. DOI:10.3389/fphys.2013.00412.

46. Donsante, A., Miller, D. G., Li, Y., Vogler, C., Brunt, E. M., Russell, D. W. and Sands, M. S. (2007) AAV Vector Integration Sites in Mouse Hepatocellular Carcinoma, *Science*, 317 (5837), pp. 477–477. DOI:10.1126/science.1142658.

47. Farr, G. A., Cotmore, S. F. and Tattersall, P. (2006) VP2 cleavage and the leucine ring at the base of the fivefold cylinder control pH-dependent externalization of both the VP1 N terminus and the genome of minute virus of mice., *Journal of Virology*, 80 (1), pp. 161–71. DOI:10.1128/JVI.80.1.161-171.2006.

48. Farr, G. A. and Tattersall, P. (2004) A conserved leucine that constricts the pore through the capsid fivefold cylinder plays a central role in parvoviral infection., *Virology*, 323 (2), pp. 243–56. DOI:10.1016/j.virol.2004.03.006.

49. Farr, G. A., Zhang, L. and Tattersall, P. (2005) Parvoviral virions deploy a capsid-tethered lipolytic enzyme to breach the endosomal membrane during cell entry., *Proceedings of the National Academy of Sciences of the United States of America*, 102 (47), pp. 17148–53. DOI:10.1073/pnas.0508477102.

50. Fay, N. and Panté, N. (2015) Nuclear entry of DNA viruses, *Frontiers in Microbiology*, 6 (MAY), pp. 1–19. DOI:10.3389/fmicb.2015.00467.

51. Feschotte, C. and Gilbert, C. (2012) Endogenous viruses: insights into viral evolution and impact on host biology, *Nature Reviews Genetics*, 13 (4), pp. 283–296. DOI:10.1038/nrg3199.

52. Flint, S. J., Enquist, L. W. and Racaniello, R. V. (2008) *Principles of virology*. ASM Press.

References

53. Gao, G.-P., Alvira, M. R., Wang, L., Calcedo, R., Johnston, J. and Wilson, J. M. (2002) Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy, *Proceedings of the National Academy of Sciences*, 99 (18), pp. 11854–11859. DOI:10.1073/pnas.182412299.

54. Gao, G., Alvira, M. R., Somanathan, S., Lu, Y., Vandenberghe, L. H., Rux, J. J., *et al.* (2003) Adeno-associated viruses undergo substantial evolution in primates during natural infections., *Proceedings of the National Academy of Sciences of the United States of America*, 100 (10), pp. 6081–6. DOI:10.1073/pnas.0937739100.

55. Gao, G., Vandenberghe, L. H., Alvira, M. R., Lu, Y., Calcedo, R., Zhou, X. and Wilson, J. M. (2004) Clades of Adeno-associated viruses are widely disseminated in human tissues., *Journal of Virology*, 78 (12), pp. 6381–8. DOI:10.1128/JVI.78.12.6381-6388.2004.

56. Garcin, P. O. and Panté, N. (2015) The minute virus of mice exploits different endocytic pathways for cellular uptake., *Virology*, 482, pp. 157–66. DOI:10.1016/j.virol.2015.02.054.

57. Gardiner, E. M. and Tattersall, P. (1988) Mapping of the fibrotropic and lymphotropic host range determinants of the parvovirus minute virus of mice., *Journal of Virology*, 62 (8), pp. 2605–13.

58. Gil-Ranedo, J., Hernando, E., Riolobos, L., Domínguez, C., Kann, M. and Almendral, J. M. (2015) The Mammalian Cell Cycle Regulates Parvovirus Nuclear Capsid Assembly., *PLoS Pathogens*, 11 (6), pp. e1004920. DOI:10.1371/journal.ppat.1004920.

59. Gil-Ranedo, J., Hernando, E., Valle, N., Riolobos, L., Maroto, B. and Almendral, J. M. (2018) Differential phosphorylation and n-terminal configuration of capsid subunits in parvovirus assembly and viral trafficking, *Virology*, 518, pp. 184–194. DOI:10.1016/j.virol.2018.02.018.

60. Guerra, S. (2000) *Interacciones del parvovirus MVM con células transformadas humanas: Regulación de la replicación viral por fosforilación de la proteína NS-1*. Autonomous University of Madrid, Spain.

61. Guindon, S. and Gascuel, O. (2003) A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood, *Systematic Biology*, 52 (5), pp. 696–704. DOI:10.1080/10635150390235520.

62. Gurda, B. L., Parent, K. N., Bladek, H., Sinkovits, R. S., DiMattia, M. A., Rence, C., *et al.* (2010) Human bocavirus capsid structure: insights into the structural repertoire of the parvoviridae., *Journal of Virology*, 84 (12), pp. 5880–9. DOI:10.1128/JVI.02719-09.

63. Halder, S., Cotmore, S., Heimburg-Molinaro, J., Smith, D. F., Cummings, R. D., Chen, X., *et al.* (2014) Profiling of Glycan Receptors for Minute Virus of Mice in Permissive Cell Lines Towards Understanding the Mechanism of Cell Recognition, *PLoS ONE*, 9 (1), pp. e86909. DOI:10.1371/journal.pone.0086909.

64. Hamilton, H., Gomos, J., Berns, K. I. and Falck-Pedersen, E. (2004) Adeno-associated virus site-specific integration and AAVS1 disruption., *Journal of Virology*, 78 (15), pp. 7874–82. DOI:10.1128/JVI.78.15.7874-7882.2004.

65. Harbison, C. E., Chiorini, J. A. and Parrish, C. R. (2008) The parvovirus capsid odyssey: from the cell surface to the nucleus, *Trends in Microbiology*, 16 (5), pp. 208–214. DOI:10.1016/j.tim.2008.01.012.

66. Hendrie, P. C., Hirata, R. K. and Russell, D. W. (2003) Chromosomal integration and homologous gene targeting by replication-incompetent vectors based on the autonomous parvovirus minute virus of mice., *Journal of Virology*, 77 (24), pp. 13136–45. DOI:10.1128/JVI.77.24.13136-13145.2003.

References

67. Hensens, O. D., Monaghan, R. L., Huang, L. and Albers-Schonberg, G. (1983) Structure of the sodium and potassium ion activated adenosine triphosphatase inhibitor L-681,110, *Journal of the American Chemical Society*, 105 (11), pp. 3672–3679. DOI:10.1021/ja00349a054.

68. Hernando, E., Llamas-Saiz, A. L., Foces-Foces, C., McKenna, R., Portman, I., Agbandje-McKenna, M. and Almendral, J. M. (2000) Biochemical and Physical Characterization of Parvovirus Minute Virus of Mice Virus-like Particles, *Virology*, 267 (2), pp. 299–309. DOI:10.1006/viro.1999.0123.

69. Hirt, B., Colomar, M. C. and Beard, P. (1998) Two segments in the genome of the immunosuppressive minute virus of mice determine the host-cell specificity, control viral DNA replication and affect viral RNA metabolism., *Journal of General Virology*, 79 (3), pp. 581–586. DOI:10.1099/0022-1317-79-3-581.

70. Hogle, J. M. (2002) Poliovirus Cell Entry: Common Structural Themes in Viral Cell Entry Pathways, *Annual Review of Microbiology*, 56 (1), pp. 677–702. DOI:10.1146/annurev.micro.56.012302.160757.

71. Horie, M., Honda, T., Suzuki, Y., Kobayashi, Y., Daito, T., Oshida, T., *et al.* (2010) Endogenous non-retroviral RNA virus elements in mammalian genomes, *Nature*, 463 (7277), pp. 84–87. DOI:10.1038/nature08695.

72. Hörlein, R., Stremmel, W., Rommelaere, J., Moroianu, A. and Willwand, K. (2002) Specific interaction of the nonstructural protein NS1 of minute virus of mice (MVM) with [ACCA]2 motifs in the centre of the right-end MVM DNA palindrome induces hairpin-primed viral DNA replication, *Journal of General Virology*, 83 (7), pp. 1659–1664. DOI:10.1099/0022-1317-83-7-1659.

73. Hu, Y. B., Dammer, E. B., Ren, R. J. and Wang, G. (2015) The endosomal-lysosomal system: From acidification and cargo sorting to neurodegeneration, *Translational Neurodegeneration*. BioMed Central. DOI:10.1186/s40035-015-0041-1.

74. Hueffer, K., Palermo, L. M. and Parrish, C. R. (2004) Parvovirus infection of cells by using variants of the feline transferrin receptor altering clathrin-mediated endocytosis, membrane domain localization, and capsid-binding domains., *Journal of Virology*, 78 (11), pp. 5601–11. DOI:10.1128/JVI.78.11.5601-5611.2004.

75. Hueffer, K., Parker, J. S. L., Weichert, W. S., Geisel, R. E., Sgro, J.-Y. and Parrish, C. R. (2003) The natural host range shift and subsequent evolution of canine parvovirus resulted from virus-specific binding to the canine transferrin receptor., *Journal of Virology*, 77 (3), pp. 1718–26. DOI:10.1128/JVI.77.3.1718-1726.2003.

76. Hüser, D., Gogol-Döring, A., Lutter, T., Weger, S., Winter, K., Hammer, E.-M., Cathomen, T., Reinert, K. and Heilbronn, R. (2010) Integration Preferences of Wildtype AAV-2 for Consensus Rep-Binding Sites at Numerous Loci in the Human Genome, *PLoS Pathogens*, 6 (7), pp. e1000985. DOI:10.1371/journal.ppat.1000985.

77. Hutagalung, A. H. and Novick, P. J. (2011) Role of Rab GTPases in membrane traffic and cell physiology., *Physiological Reviews*, 91 (1), pp. 119–49. DOI:10.1152/physrev.00059.2009.

78. Jagdale, S. S. and Joshi, R. S. (2018) Enemies with benefits: mutualistic interactions of viruses with lower eukaryotes, *Archives of Virology*, 163 (4), pp. 821–830. DOI:10.1007/s00705-017-3686-5.

79. Janovitz, T., Klein, I. A., Oliveira, T., Mukherjee, P., Nussenzweig, M. C., Sadelain, M. and Falck-Pedersen, E. (2013) High-Throughput Sequencing Reveals Principles of Adeno-Associated Virus Serotype 2 Integration, *Journal of Virology*, 87 (15), pp. 8559–8568. DOI:10.1128/JVI.01135-13.

80. Janovitz, T., Oliveira, T., Sadelain, M. and Falck-Pedersen, E. (2014) Highly divergent integration profile of adeno-associated virus serotype 5 revealed by high-throughput sequencing., *Journal of Virology*, 88 (5), pp. 2481–8. DOI:10.1128/JVI.03419-13.

References

81. Jin, H., Xia, X., Liu, B., Fu, Y., Chen, X., Wang, H. and Xia, Z. (2016) High-yield production of canine parvovirus virus-like particles in a baculovirus expression system, *Archives of Virology*, 161 (3), pp. 705–710. DOI:10.1007/s00705-015-2719-1.

82. Kaelber, J. T., Demogines, A., Harbison, C. E., Allison, A. B., Goodman, L. B., Ortega, A. N., Sawyer, S. L. and Parrish, C. R. (2012) Evolutionary reconstructions of the transferrin receptor of Caniforms supports canine parvovirus being a re-emerged and not a novel pathogen in dogs., *PLoS Pathogens*, 8 (5), pp. e1002666. DOI:10.1371/journal.ppat.1002666.

83. Kailasan, S., Agbandje-McKenna, M. and Parrish, C. R. (2015) Parvovirus Family Conundrum: What Makes a Killer?, *Annual Review of Virology*, 2 (1), pp. 425–450. DOI:10.1146/annurev-virology-100114-055150.

84. Kapoor, A., Simmonds, P. and Lipkin, W. I. (2010) Discovery and characterization of mammalian endogenous parvoviruses., *Journal of Virology*, 84 (24), pp. 12628–35. DOI:10.1128/JVI.01732-10.

85. Katzourakis, A. and Gifford, R. J. (2010) Endogenous viral elements in animal genomes., *PLoS Genetics*, 6 (11), pp. e1001191. DOI:10.1371/journal.pgen.1001191.

86. Kaufmann, B., López-Bueno, A., Mateu, M. G., Chipman, P. R., Nelson, C. D. S., Parrish, C. R., Almendral, J. M. and Rossmann, M. G. (2007) Minute virus of mice, a parvovirus, in complex with the Fab fragment of a neutralizing monoclonal antibody., *Journal of Virology*, 81 (18), pp. 9851–8. DOI:10.1128/JVI.00775-07.

87. Kimsey, P. B., Engers, H. D., Hirt, B. and Jongeneel, C. V (1986) Pathogenicity of fibroblast- and lymphocyte-specific variants of minute virus of mice., *Journal of Virology*, 59 (1), pp. 8–13.

88. Kontou, M., Govindasamy, L., Nam, H.-J. H.-J., Bryant, N., Llamas-Saiz, A. L., Foces-Foces, C., *et al.* (2005) Structural determinants of tissue tropism and in vivo pathogenicity for the parvovirus minute virus of mice., *Journal of Virology*, 79 (17), pp. 10931–43. DOI:10.1128/JVI.79.17.10931-10943.2005.

89. Kotin, R. M., Linden, R. M. and Berns, K. I. (1992) Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination., *The EMBO Journal*, 11 (13), pp. 5071–8. DOI:10.1182/blood-2008-10-181479.

90. Krupovic, M. (2013) Networks of evolutionary interactions underlying the polyphyletic origin of ssDNA viruses, *Current Opinion in Virology*, 3 (5), pp. 578–586. DOI:10.1016/J.COVIRO.2013.06.010.

91. Linser, P., Bruning, H. and Armentrout, R. W. (1977) Specific binding sites for a parvovirus, minute virus of mice, on cultured mouse cells., *Journal of Virology*, 24 (1), pp. 211–21

92. Liu, H., Fu, Y., Xie, J., Cheng, J., Ghabrial, S. A., Li, G., Peng, Y., Yi, X. and Jiang, D. (2011) Widespread endogenization of densoviruses and parvoviruses in animal and human genomes., *Journal of Virology*, 85 (19), pp. 9863–9876. DOI:10.1128/JVI.00828-11.

93. Lombardo, E., Ramírez, J. C., Agbandje-McKenna, M. and Almendral, J. M. (2000) A beta-stranded motif drives capsid protein oligomers of the parvovirus minute virus of mice into the nucleus for viral assembly., *Journal of Virology*, 74 (8), pp. 3804–14. DOI:10.1128/JVI.74.8.3804-3814.2000.

94. Lombardo, E., Ramírez, J. C., Garcia, J. and Almendral, J. M. (2002) Complementary roles of multiple nuclear targeting signals in the capsid proteins of the parvovirus minute virus of mice during assembly and onset of infection., *Journal of Virology*, 76 (14), pp. 7049–59. DOI:10.1128/JVI.76.14.7049-7059.2002.

References

95. López-Bueno, A., Mateu, M. G. and Almendral, J. M. (2003) High mutant frequency in populations of a DNA virus allows evasion from antibody therapy in an immunodeficient host., *Journal of Virology*, 77 (4), pp. 2701–8. DOI:10.1128/JVI.77.4.2701-2708.2003.

96. Lopez-Bueno, A., Rubio, M.-P., Bryant, N., McKenna, R., Agbandje-McKenna, M. and Almendral, J. M. (2006) Host-Selected Amino Acid Changes at the Sialic Acid Binding Pocket of the Parvovirus Capsid Modulate Cell Binding Affinity and Determine Virulence, *Journal of Virology*, 80 (3), pp. 1563–1573. DOI:10.1128/JVI.80.3.1563-1573.2006.

97. Lopez-Bueno, A., Segovia, J. C., Bueren, J. A., O'Sullivan, M. G., Wang, F., Tattersall, P. and Almendral, J. M. (2008) Evolution to Pathogenicity of the Parvovirus Minute Virus of Mice in Immunodeficient Mice Involves Genetic Heterogeneity at the Capsid Domain That Determines Tropism, *Journal of Virology*, 82 (3), pp. 1195–1203. DOI:10.1128/JVI.01692-07.

98. López-Bueno, A., Villarreal, L. P. and Almendral, J. M. (2006) Parvovirus Variation for Disease: A Difference with RNA Viruses?, in: *Quasispecies: Concept and Implications for Virology*. Berlin/Heidelberg: Springer-Verlag,299, pp. 349–370.

99. Luisoni, S. and Greber, U. F. (2016) Biology of Adenovirus Cell Entry, in: *Adenoviral Vectors for Gene Therapy*. Elsevier, pp. 27–58.

100. Maggin, J. E., James, J. A., Chappie, J. S., Dyda, F. and Hickman, A. B. (2012) The Amino Acid Linker between the Endonuclease and Helicase Domains of Adeno-Associated Virus Type 5 Rep Plays a Critical Role in DNA-Dependent Oligomerization, *Journal of Virology*, 86 (6), pp. 3337–3346. DOI:10.1128/JVI.06775-11.

101. Mani, B., Baltzer, C., Valle, N., Almendral, J. M., Kempf, C. and Ros, C. (2006) Low pH-dependent endosomal processing of the incoming parvovirus minute virus of mice virion leads to externalization of the VP1 N-terminal sequence (N-VP1), N-VP2 cleavage, and uncoating of the full-length genome., *Journal of Virology*, 80 (2), pp. 1015–24. DOI:10.1128/JVI.80.2.1015-1024.2006.

102. Marchini, A., Bonifati, S., Scott, E. M., Angelova, A. L. and Rommelaere, J. (2015) Oncolytic parvoviruses: from basic virology to clinical applications, *Virology Journal*, 12 (1), pp. 6. DOI:10.1186/s12985-014-0223-y.

103. Maroto, B., Ramírez, J. C. and Almendral, J. M. (2000) Phosphorylation status of the parvovirus minute virus of mice particle: mapping and biological relevance of the major phosphorylation sites., *Journal of Virology*, 74 (23), pp. 10892–902. DOI:10.1128/JVI.74.23.10892-10902.2000.

104. Maroto, B., Valle, N., Saffrich, R. and Almendral, J. M. (2004) Nuclear Export of the Nonenveloped Parvovirus Virion Is Directed by an Unordered Protein Signal Exposed on the Capsid Surface, *Journal of Virology*, 78 (19), pp. 10685–10694. DOI:10.1128/JVI.78.19.10685-10694.2004.

105. Maxwell, I. H., Spitzer, A. L., Maxwell, F. and Pintel, D. J. (1995) The capsid determinant of fibrotropism for the MVMp strain of minute virus of mice functions via VP2 and not VP1., *Journal of Virology*, 69 (9), pp. 5829–32.

106. Medrano, M., Fuertes, M. Á., Valbuena, A., Carrillo, P. J. P., Rodríguez-Huete, A. and Mateu, M. G. (2016) Imaging and Quantitation of a Succession of Transient Intermediates Reveal the Reversible Self-Assembly Pathway of a Simple Icosahedral Virus Capsid, *Journal of the American Chemical Society*, 138 (47), pp. 15385–15396. DOI:10.1021/jacs.6b07663.

107. Miller, C. L. C. and Pintel, D. D. J. (2002) Interaction between parvovirus NS2 protein and nuclear export factor Crm1 is important for viral egress from the nucleus of murine cells., *Journal of Virology*, 76 (7), pp. 3257–66. DOI:10.1128/JVI.76.7.3257.

108. Mouw, M. and Pintel, D. J. (1998) Amino acids 16-275 of minute virus of mice NS1 include a domain that specifically binds (ACCA)2-3-containing DNA, *Virology*, 251 (1), pp. 123–131. DOI:10.1006/viro.1998.9375.

References

109. Naccache, S. N., Greninger, A. L., Lee, D., Coffey, L. L., Phan, T., Rein-Weston, A., *et al.* (2013) The Perils of Pathogen Discovery: Origin of a Novel Parvovirus-Like Hybrid Genome Traced to Nucleic Acid Extraction Spin Columns, *Journal of Virology*, 87 (22), pp. 11966–11977. DOI:10.1128/JVI.02323-13.

110. Naeger, L. K., Cater, J. and Pintel, D. J. (1990) The small nonstructural protein (NS2) of the parvovirus minute virus of mice is required for efficient DNA replication and infectious virus production in a cell-type-specific manner., *Journal of Virology*, 64 (12), pp. 6166–6175.

111. Naeger, L. K., Salomé, N. and Pintel, D. J. (1993) NS2 is required for efficient translation of viral mRNA in minute virus of mice-infected murine cells., *Journal of Virology*, 67 (2), pp. 1034–43.

112. Nam, H.-J., Gurda-Whitaker, B., Gan, W. Y., Ilaria, S., McKenna, R., Mehta, P., Alvarez, R. A. and Agbandje-McKenna, M. (2006) Identification of the sialic acid structures recognized by minute virus of mice and the role of binding affinity in virulence adaptation., *The Journal of Biological Chemistry*, 281 (35), pp. 25670–7. DOI:10.1074/jbc.M604421200.

113. Nault, J.-C., Datta, S., Imbeaud, S., Franconi, A., Mallet, M., Couchy, G., *et al.* (2015) Recurrent AAV2-related insertional mutagenesis in human hepatocellular carcinomas, *Nature Genetics*, 47 (10), pp. 1187–1193. DOI:10.1038/ng.3389.

114. Nüesch, J. P. F., Cotmore, S. F. and Tattersall, P. (1995) Sequence motifs in the replicator protein of parvovirus MVM essential for nicking and covalent attachment to the viral origin: Identification of the linking tyrosine, *Virology*, 209 (1), pp. 122–135. DOI:10.1006/viro.1995.1236.

115. Nüesch, J. P. F., Lacroix, J., Marchini, A. and Rommelaere, J. (2012) Molecular pathways: Rodent parvoviruses - Mechanisms of oncolysis and prospects for clinical cancer treatment, *Clinical Cancer Research*, 18 (13), pp. 3516–3523. DOI:10.1158/1078-0432.CCR-11-2325.

116. Nüesch, J. P. F. and Tattersall, P. (1993) Nuclear targeting of the parvoviral replicator molecule NS1: Evidence for self-association prior to nuclear transport, *Virology*, 196 (2), pp. 637–651. DOI:10.1006/viro.1993.1520.

117. Ohkuma, S. and Poole, B. (1978) Fluorescence probe measurement of the intralysosomal pH in living cells and the perturbation of pH by various agents., *Proceedings of the National Academy of Sciences of the United States of America*, 75 (7), pp. 3327–31. DOI:10.1073/pnas.75.7.3327.

118. Olofsson, S. and Bergström, T. (2005) Glycoconjugate glycans as viral receptors, *Annals of Medicine*, 37 (3), pp. 154–172. DOI:10.1080/07853890510007340.

119. Palermo, L. M., Hafenstein, S. L. and Parrish, C. R. (2006) Purified feline and canine transferrin receptors reveal complex interactions with the capsids of canine and feline parvoviruses that correspond to their host ranges., *Journal of Virology*, 80 (17), pp. 8482–92. DOI:10.1128/JVI.00683-06.

120. Palermo, L. M., Hueffer, K. and Parrish, C. R. (2003) Residues in the apical domain of the feline and canine transferrin receptors control host-specific binding and cell infection of canine and feline parvoviruses., *Journal of Virology*, 77 (16), pp. 8915–23. DOI:10.1128/JVI.77.16.8915-8923.2003.

121. Paradiso, P. R. (1981) Infectious process of the parvovirus H-1: correlation of protein content, particle density, and viral infectivity., *Journal of Virology*, 39 (3), pp. 800–7.

122. Paradiso, P. R., Williams, K. R. and Costantino, R. L. (1984) Mapping of the amino terminus of the H-1 parvovirus major capsid protein., *Journal of Virology*, 52 (1), pp. 77–81.

References

123. Parker, J. S. L., Murphy, W. J., Wang, D., O'Brien, S. J. and Parrish, C. R. (2001) Canine and feline parvoviruses can use human or feline transferrin receptors to bind, enter, and infect cells., *Journal of Virology*, 75 (8), pp. 3896–902. DOI:10.1128/JVI.75.8.3896-3902.2001.

124. Parrish, C. R. (2010) Structures and Functions of Parvovirus Capsids and the Process of Cell Infection, *Current Topics in Microbiology and Immunology*, 343, pp. 149–176. DOI:10.1007/82_2010_33.

125. Perelman, P., Johnson, W. E., Roos, C., Seuánez, H. N., Horvath, J. E., Moreira, M. A. M., *et al.* (2011) A Molecular Phylogeny of Living Primates, *PLoS Genetics*, 7 (3), pp. e1001342. DOI:10.1371/journal.pgen.1001342.

126. Petter, J.-J. and Desbordes, F. (2010) *Primates of the world : an illustrated guide*. Princeton University Press, New Jersey (USA) and Oxfordshire (UK).

127. Qiu, J., Söderlund-Venermo, M. and Young, N. S. Human parvoviruses, , 30 Clinical Microbiology Reviews 43–113 (2017). DOI:10.1128/CMR.00040-16.

128. Ramírez, J. C., Fairén, A. and Almendral, J. M. (1996) Parvovirus minute virus of mice strain i multiplication and pathogenesis in the newborn mouse brain are restricted to proliferative areas and to migratory cerebellar young neurons., *Journal of Virology*, 70 (11), pp. 8109–16.

129. Ramírez, J. C., Santarén, J. F. and Almendral, J. M. (1995) Transcriptional inhibition of the parvovirus minute virus of mice by constitutive expression of an antisense RNA targeted against the NS-1 transactivator protein, *Virology*, 206 (1), pp. 57–68. DOI:10.1016/S0042-6822(95)80019-0.

130. Ran, F. A., Cong, L., Yan, W. X., Scott, D. A., Gootenberg, J. S., Kriz, A. J., *et al.* (2015) In vivo genome editing using Staphylococcus aureus Cas9, *Nature*, 520 (7546), pp. 186–191. DOI:10.1038/nature14299.

References

131. Reguera, J., Carreira, A., Riolobos, L., Almendral, J. M. and Mateu, M. G. (2004) Role of interfacial amino acid residues in assembly, stability, and conformation of a spherical virus capsid, *Proceedings of the National Academy of Sciences*, 101 (9), pp. 2724–2729. DOI:10.1073/pnas.0307748101.

132. Riolobos, L., Valle, N., Hernando, E., Maroto, B., Kann, M. and Almendral, J. M. (2010) Viral oncolysis that targets Raf-1 signaling control of nuclear transport., *Journal of Virology*, 84 (4), pp. 2090–9. DOI:10.1128/JVI.01550-09.

133. Roossinck, M. J. and Bazán, E. R. (2017) Symbiosis: Viruses as Intimate Partners, *Annual Review of Virology*, 4 (1), pp. 123–139. DOI:10.1146/annurev-virology-110615-042323.

134. Ros, C., Baltzer, C., Mani, B. and Kempf, C. (2006) Parvovirus uncoating in vitro reveals a mechanism of DNA release without capsid disassembly and striking differences in encapsidated DNA stability, *Virology*, 345 (1), pp. 137–147. DOI:10.1016/j.virol.2005.09.030.

135. Ros, C., Bayat, N., Wolfisberg, R. and Almendral, J. (2017) Protoparvovirus Cell Entry, *Viruses*, 9 (11), pp. 313. DOI:10.3390/v9110313.

136. Ros, C., Burckhardt, C. J. and Kempf, C. (2002) Cytoplasmic trafficking of minute virus of mice: low-pH requirement, routing to late endosomes, and proteasome interaction., *Journal of Virology*, 76 (24), pp. 12634–12645. DOI:10.1128/JVI.76.24.12634-12645.2002.

137. Rossmann, M. G. and Johnson, J. E. (1989) Icosahedral RNA Virus Structure, *Annual Review of Biochemistry*, 58 (1), pp. 533–569. DOI:10.1146/annurev.bi.58.070189.002533.

138. Rubio, M.-P., López-Bueno, A. and Almendral, J. M. (2005) Virulent variants emerging in mice infected with the apathogenic prototype strain of the parvovirus minute virus of mice exhibit a capsid with low avidity for a primary receptor., *Journal of Virology*, 79 (17), pp. 11280–90. DOI:10.1128/JVI.79.17.11280-11290.2005.

References

139. Rubio, M. P., Guerra, S. and Almendral, J. M. (2001) Genome replication and postencapsidation functions mapping to the nonstructural gene restrict the host range of a murine parvovirus in human cells., *Journal of Virology*, 75 (23), pp. 11573–82. DOI:10.1128/JVI.75.23.11573-11582.2001.

140. Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees., *Molecular Biology and Evolution*, 4 (4), pp. 406–25. DOI:10.1093/oxfordjournals.molbev.a040454.

141. Samulski, R. J., Zhu, X., Xiao, X., Brook, J. D., Housman, D. E., Epstein, N. and Hunter, L. A. (1991) Targeted integration of adeno-associated virus (AAV) into human chromosome 19., *The EMBO Journal*, 10 (12), pp. 3941–50. DOI:10.1523/JNEUROSCI.4323-08.2008.

142. Sánchez-Martínez, C., Grueso, E., Carroll, M., Rommelaere, J. and Almendral, J. M. (2012) Essential role of the unordered VP2 n-terminal domain of the parvovirus MVM capsid in nuclear assembly and endosomal enlargement of the virion fivefold channel for cell entry., *Virology*, 432 (1), pp. 45–56. DOI:10.1016/j.virol.2012.05.025.

143. Santarén, J. F., Ramírez, J. C. and Almendral, J. M. (1993) Protein species of the parvovirus minute virus of mice strain MVMp: involvement of phosphorylated VP-2 subtypes in viral morphogenesis., *Journal of Virology*, 67 (9), pp. 5126–38.

144. Schnepp, B. C., Jensen, R. L., Chen, C.-L., Johnson, P. R. and Clark, K. R. (2005) Characterization of adeno-associated virus genomes isolated from human tissues., *Journal of Virology*, 79 (23), pp. 14793–803. DOI:10.1128/JVI.79.23.14793-14803.2005.

145. Segovia, J. C., Gallego, J. M., Bueren, J. A. and Almendral, J. M. (1999) Severe leukopenia and dysregulated erythropoiesis in SCID mice persistently infected with the parvovirus minute virus of mice., *Journal of Virology*, 73 (3), pp. 1774–84.

146. Segovia, J. C., Guenechea, G., Gallego, J. M., Almendral, J. M. and Bueren, J. A. (2003) Parvovirus infection suppresses long-term repopulating hematopoietic stem cells., *Journal of Virology*, 77 (15), pp. 8495–503. DOI:10.1128/JVI.77.15.8495-8503.2003.

147. Segovia, J. C., Real, A., Bueren, J. A. and Almendral, J. M. (1991) In vitro myelosuppressive effects of the parvovirus minute virus of mice (MVMi) on hematopoietic stem and committed progenitor cells., *Blood*, 77 (5), pp. 980–8.

148. Shein, H. and Eneders, J. (1962) Multiplication and cytopathogenicity of Simian vacuolating virus 40 in cultures of human tissues, *Proc Soc Enp Biol Med*, 109, pp. 495–500. DOI:10.3181/00379727-109-27246.

149. Shein, H. M., Enders, J. F. and Levinthal, J. D. (1962) Transformation induced by simian virus 40 in human renal cell cultures. II. Cell-virus relationships., *Proceedings of the National Academy of Sciences of the United States of America*, 48, pp. 1350–7. DOI:10.1073/pnas.48.8.1350.

150. Sievers, F. and Higgins, D. G. (2014) Clustal Omega, in: *Current Protocols in Bioinformatics*. Hoboken, NJ, USA: John Wiley & Sons, Inc.,48, pp. 3.13.1-3.13.16.

151. Smuts, H., Kew, M., Khan, A. and Korsman, S. (2014) Novel hybrid parvovirus-like virus, NIH-CQV/PHV, contaminants in silica column-based nucleic acid extraction kits., *Journal of Virology*, 88 (2), pp. 1398. DOI:10.1128/JVI.03206-13.

152. Spalholz, B. A. and Tattersall, P. (1983) Interaction of minute virus of mice with differentiated cells: strain-dependent target cell specificity is mediated by intracellular factors., *Journal of Virology*, 46 (3), pp. 937–43.

153. States, D. J. and Gish, W. (1994) Combined use of sequence similarity and codon bias for coding region identification., *Journal of Computational Biology : A Journal of Computational Molecular Cell Biology*, 1 (1), pp. 39–50. DOI:10.1089/cmb.1994.1.39.

References

154. Stutika, C., Gogol-Döring, A., Botschen, L., Mietzsch, M., Weger, S., Feldkamp, M., Chen, W. and Heilbronn, R. (2016) A Comprehensive RNA Sequencing Analysis of the Adeno-Associated Virus (AAV) Type 2 Transcriptome Reveals Novel AAV Transcripts, Splice Variants, and Derived Proteins., *Journal of Virology*, 90 (3), pp. 1278–89. DOI:10.1128/JVI.02750-15.

155. Subramanian, S., Organtini, L. J., Grossman, A., Domeier, P. P., Cifuente, J. O., Makhov, A. M., *et al.* (2017) Cryo-EM maps reveal five-fold channel structures and their modification by gatekeeper mutations in the parvovirus minute virus of mice (MVM) capsid, *Virology*, 510, pp. 216–223. DOI:10.1016/j.virol.2017.07.015.

156. Suikkanen, S., Aaltonen, T., Nevalainen, M., Välilehto, O., Lindholm, L., Vuento, M. and Vihinen-Ranta, M. (2003) Exploitation of microtubule cytoskeleton and dynein during parvoviral traffic toward the nucleus., *Journal of Virology*, 77 (19), pp. 10270–9. DOI:10.1128/JVI.77.19.10270-10279.2003.

157. Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0, *Molecular Biology and Evolution*, 30 (12), pp. 2725–2729. DOI:10.1093/molbev/mst197.

158. Tarailo-Graovac, M. and Chen, N. (2009) Using RepeatMasker to Identify Repetitive Elements in Genomic Sequences, in: *Current Protocols in Bioinformatics*. Hoboken, NJ, USA: John Wiley & Sons, Inc.,Chapter 4, pp. Unit 4.10.

159. Tattersall, P. and Bratton, J. (1983) Reciprocal productive and restrictive virus-cell interactions of immunosuppressive and prototype strains of minute virus of mice., *Journal of Virology*, 46 (3), pp. 944–55.

160. Tattersall, P., Shatkin, A. J. and Ward, D. C. (1977) Sequence homology between the structural polypeptides of minute virus of mice, *Journal of Molecular Biology*, 111 (4), pp. 375–394. DOI:10.1016/S0022-2836(77)80060-0.

References

161. Tattersall, P. and Ward, D. C. (1976) Rolling hairpin model for replication of parvovirus and linear chromosomal DNA, *Nature*, 263 (5573), pp. 106–109. DOI:10.1038/263106a0.

162. Taylor, D. J., Leach, R. W. and Bruenn, J. (2010) Filoviruses are ancient and integrated into mammalian genomes, *BMC Evolutionary Biology*, 10 (1), pp. 193. DOI:10.1186/1471-2148-10-193.

163. Tijssen, P., Agbandje-McKenna, M., Almendral, J. M., Bergoin, M., Flegel, T. W., Hedman, K., *et al.* (2012) Parvoviridae, in: King AMQ, Adams MJ, Carstens EB, L. E. (ed.) *Virus Taxonomy*. London: Elsevier, pp. 405–425.

164. Timney, B. L., Raveh, B., Mironska, R., Trivedi, J. M., Kim, S. J., Russel, D., Wente, S. R., Sali, A. and Rout, M. P. (2016) Simple rules for passive diffusion through the nuclear pore complex, *The Journal of Cell Biology*, 215 (1), pp. 57–76. DOI:10.1083/jcb.201601004.

165. Trempe, J. P. and Carter, B. J. (1988) Alternate mRNA splicing is required for synthesis of adeno-associated virus VP1 capsid protein., *Journal of Virology*, 62 (9), pp. 3356–3363.

166. Tsai, B. and Qian, M. (2010) Cellular Entry of Polyomaviruses, in: *Current topics in microbiology and immunology*.343, pp. 177–194.

167. Tsao, J., Chapman, M. S., Agbandje, M., Keller, W., Smith, K., Wu, H., *et al.* (1991) The three-dimensional structure of canine parvovirus and its functional implications., *Science (New York, N.Y.)*, 251 (5000), pp. 1456–64. DOI:10.2210/pdb2dpv/pdb.

168. Tu, M., Liu, F., Chen, S., Wang, M. and Cheng, A. (2015) Role of capsid proteins in parvoviruses infection, *Virology Journal*, 12 (1), pp. 114. DOI:10.1186/s12985-015-0344-y.

References

169. Tullis, G. E., Burger, L. R. and Pintel, D. J. (1992) The trypsin-sensitive RVER domain in the capsid proteins of minute virus of mice is required for efficient cell binding and viral infection but not for proteolytic processing in vivo., *Virology*, 191 (2), pp. 846–57. DOI:10.1016/0042-6822(92)90260-v.

170. Tullis, G. E., Labieniec-Pintel, L., Clemens, K. E. and Pintel, D. (1988) Generation and characterization of a temperature-sensitive mutation in the NS-1 gene of the autonomous parvovirus minute virus of mice., *Journal of Virology*, 62 (8), pp. 2736–44.

171. Ventoso, I., Berlanga, J. J. and Almendral, J. M. (2010) Translation Control by Protein Kinase R Restricts Minute Virus of Mice Infection: Role in Parvovirus Oncolysis, *Journal of Virology*, 84 (10), pp. 5043–5051. DOI:10.1128/JVI.02188-09.

172. Vihinen-Ranta, M., Kakkola, L., Kalela, A., Vilja, P. and Vuento, M. (1997) Characterization of a nuclear localization signal of canine parvovirus capsid proteins., *European Journal of Biochemistry*, 250 (2), pp. 389–94. DOI:10.1111/j.1432-1033.1997.0389a.x.

173. Vihinen-Ranta, M., Kalela, A., Mäkinen, P., Kakkola, L., Marjomäki, V. and Vuento, M. (1998) Intracellular route of canine parvovirus entry., *Journal of Virology*, 72 (1), pp. 802–6.

174. Vihinen-Ranta, M., Suikkanen, S. and Parrish, C. R. (2004) Pathways of cell infection by parvoviruses and adeno-associated viruses., *Journal of Virology*, 78 (13), pp. 6709–14. DOI:10.1128/JVI.78.13.6709-6714.2004.

175. Vihinen-Ranta, M., Wang, D., Weichert, W. S. and Parrish, C. R. (2002) The VP1 N-terminal sequence of canine parvovirus affects nuclear transport of capsids and efficient cell infection., *Journal of Virology*, 76 (4), pp. 1884–91. DOI:10.1128/JVI.76.4.1884-1891.2002.

176. Walker, S. L., Wonderling, R. S. and Owens, R. A. (1997) Mutational analysis of the adeno-associated virus type 2 Rep68 protein helicase motifs., *Journal of Virology*, 71 (9), pp. 6996–7004.

177. Weiss, R. A. and Stoye, J. P. (2013) Virology. Our viral inheritance., *Science (New York, N.Y.)*, 340 (6134), pp. 820–1. DOI:10.1126/science.1235148.

178. Weitzman MD (2006) The parvovirus life cycle: an introduction to molecular interactions important for infection., in: Kerr JR, Cotmore SF, Bloom ME, Linden RM, P. C. (ed.) *Parvoviruses*. London: Hodder Arnold, pp. 143–156.

179. Wente, S. R. and Rout, M. P. (2010) The Nuclear Pore Complex and Nuclear Transport, *Cold Spring Harbor Perspectives in Biology*, 2 (10), pp. a000562–a000562. DOI:10.1101/cshperspect.a000562.

180. Wolfisberg, R., Kempf, C. and Ros, C. (2016) Late maturation steps preceding selective nuclear export and egress of progeny parvovirus., *Journal of Virology*, 90 (11), pp. 5462–5474. DOI:10.1128/JVI.02967-15.

181. Yamauchi, Y. and Helenius, A. (2013) Virus entry at a glance, *Journal of Cell Science*, 126 (6), pp. 1289–1295. DOI:10.1242/jcs.119685.

182. Yuan, W. and Parrish, C. R. (2001) Canine Parvovirus Capsid Assembly and Differences in Mammalian and Insect Cells, *Virology*, 279 (2), pp. 546–557. DOI:10.1006/viro.2000.0734.

183. Zádori, Z., Szelei, J., Lacoste, M. C., Li, Y., Gariépy, S., Raymond, P., Allaire, M., Nabi, I. R. and Tijssen, P. (2001) A viral phospholipase A2 is required for parvovirus infectivity., *Developmental Cell*, 1 (2), pp. 291–302. DOI:10.1016/S1534-5807(01)00031-4.

184. Zinkernagel, R. M., Klenerman, P. and Hengartner, H. (1997) A non-retroviral RNA virus persists in DNA form., *Nature*, 390 (6657), pp. 298–301. DOI:10.1038/36876.